PHYSIOLOGICAL MECHANISMS OF YIELD IMPROVEMENT IN HISTORICAL U.S. SOYBEAN GERMPLASM

BY

ROBERT PAUL KOESTER

DISSERTATION

Submitted in partial fulfillment of the requirements
for the degree of Doctor of Philosophy in Plant Biology
in the Graduate College of the
University of Illinois at Urbana-Champaign, 2014

Urbana, Illinois

Doctoral Committee:

Associate Professor Elizabeth A. Ainsworth, Chair
Emerson Professor Donald R. Ort
Professor Steven C. Huber
Professor Brian W. Diers

ABSTRACT

Soybean (*Glycine max* Merr.) is the world's most widely grown leguminous crop and an important source of protein and oil for food and feed. Soybean yields have increased substantially throughout the past century with yield gains widely attributed to genetic advances and improved cultivars, as well as advances in farming technology and practice. Although soybean yields do not appear to be stagnating, the current rate of gain is insufficient to meet the United Nations target of doubling crop yields by 2050. While soybean yields have been increased through traditional breeding efforts, the physiological mechanisms underlying past yield gains in the U.S. are largely unknown. Therefore, the aims of this thesis research are to gain a better understanding of the physiological basis of past improvements in soybean yield in order to help identify strategies for increasing future production.

First, in a two year experiment, twenty-four soybean cultivars released between 1923 and 2007 were grown in field trials. Physiological improvements in the efficiencies by which soybean canopies intercepted light (ε_i), converted light energy into biomass (ε_c), and partitioned biomass into seed (ε_p) were examined. Seed yield increased on average by 26.5 kg ha^{-1} yr^{-1}, and the increase in seed yield was driven by improvements in all three efficiencies. Although the time to canopy closure did not change in historical soybean cultivars, extended growing seasons and decreased lodging in more modern lines drove improvements in ε_i. Greater biomass production per unit of absorbed light resulted in improvements in ε_c. Soybean seed biomass increased at a rate greater than total above-ground biomass, resulting in an increase in ε_p. It is thought that there is little room for further improvements in ε_i and ε_p as 84 years of traditional breeding has driven these efficiencies close to theoretical maxima. ε_c is still well below its theoretical maxima and is, therefore, a target for future yield gains.

Next, in order to investigate the potential mechanisms underlying the increase in ε_c with cultivar year of release (YOR), photosynthetic (*A*) and respiratory capacity was measured within this set of historical germplasm over 3 growing seasons. Traditional soybean breeding has improved ε_c through greater rates of *A* with no change in respiratory capacity. The gains in *A* were driven by increased rates of stomatal conductance (g_s) and water use and not improved photosynthetic capacity. Thus, greater carbon gain in modern varieties was only apparent under times of ample water supply. These results suggest that as climate change increases the water

demand of crops, past strategies for increasing conversion efficiency will have reduced effectiveness.

Finally, the transcript abundance of putative "yield enhancing genes" (YEG) were determined in these historic soybean cultivars to examine potential genetic drivers of past yield advancement. YEG are single genes, that when altered, have the capacity to increase plant growth and yield. Of the fifteen YEG examined in this study, six had gene expression levels that correlated with cultivar YOR and yield in at least one growing season. Three of these genes encode Rubisco activase which is hypothesized to increase yields by increasing rates of photosynthesis. YEG also include those that encode vegetative storage proteins which are important in the storage and transfer of carbon and nitrogen to sink tissues. Future work is needed to understand the underlying mechanisms of how altered YEG transcript abundance is affecting yield and if the alteration of these genes further will lead to additional yield gains.

This dissertation research provides insight into the physiological mechanisms underlying past yield gains and identifies targets for future yield improvement. At the whole canopy level, traditional breeding efforts have increased ε_i, ε_c, and ε_p, and identified that improving ε_c has the most potential of further increasing yields. On the leaf-level, photosynthesis has been improved through greater water use, indicating that past improvements made in ε_c may not be effective in warmer future. Finally, at the genic level, putative YEG were identified as correlating with yield in soybean, and therefore, have the potential ability to increase yields further in the future.

ACKNOWLEDGEMENTS

First off, I would like to thank my Ph.D. advisor, Dr. Lisa Ainsworth, for being an excellent mentor and supporting me over the past 7 years. She encouraged me to greatness at every step, and I am fortunate to have had the opportunity to work with her. I would also like to thank my committee members, Drs. Don Ort, Steve Huber, and Brian Diers, for providing me with invaluable advice over the years. I also thank my collaborators who made this research possible and whose feedback helped me publish my work.

I thank Troy Cary, Ed Johnson, and Dr. Jeff Skoneczka for helping with planting, harvesting, experimental design, and agronomic practices. I thank Dr. Adam Davis for advising me on statistical and modelling techniques and the USDA-ARS lab group in general for listening to my talks over the years. Not only did I get practice in public speaking, but also feedback and ideas to further my research. I also thank Drs. Craig Yendrek, Kat Grennan, and Jennifer Wai for their discussions on qPCR.

It took a lot of hands to complete this work and so I thank everyone who has helped me along the way including Nick Vertin, Jackie Rodriguez, Payam Vatani, Joe Sullivan, Anna Mollineaux, Amin Ghane, Taylor Repp, Carlos Garcia, Kirsten Michon, Rachel Paul, Darshi Banan, Anna Locke, Katie Richter, Caroline Kisielinski, Matt Kemet, Matt Kendzior, and Charley Illunga. A special shout-out to Brittany Nohl who was my right arm during these projects and made working the long hours fun.

I definitely want to thank Team Ainsworth, both past and present, for their hours of gas-exchange and help over the years including Kristen Bishop, Amy Betzelberger, Justin McGrath, and Alvaro Sanz-Saez. I especially thank Kelly Gillespie for being a great undergraduate mentor and Craig Yendrek for putting up with all the graphs I've shown him. I have made a lot of great friends along the way that not only provided great advice on my research, but also knew how to kick it after work. For that, I thank Courtney Leisner, Becky Slattery, Matt Siebers, Chris Montes, Rebekah Kelly, and Wes Barber.

Finally, I want to thank my family and friends for supporting me. I especially want to thank my parents, grandparents, and two sisters for not giving me too much grief about still being in school.

TABLE OF CONTENTS

CHAPTER I: GENERAL INTRODUCTION

Soybean (*Glycine max*) is an annual, leguminous dicot that plays an important role in supplying food, feed, and fuel, for the world. Although, domesticated soybean has its origins in Northern China (Hymowitz, 1970), the United States (U.S.) is currently the largest producer of soybean accounting for 34% of world production (FAOSTAT 2013). Soybean production is the 4th greatest globally, behind corn, rice and wheat (FAOSTAT 2013). Soybean seeds are high in protein and oil, and therefore, are used in food and feed as important suppliers of key amino acids (Wilson, 2008). Further, soybean is often cultivated in rotation with corn to reduce pest populations.

Soybean yield gains in the U.S. have increased over the past century with national average increases of 23.3 kg ha^{-1} yr^{-1} (Specht *et al.*, 2014). Improvements have been achieved through the continued breeding of new cultivars, advancements in agronomic technologies, and an increase in atmospheric CO_2 (Specht *et al*., 1999; De Bruin *et al.*, 2008; Rowntree *et al.*, 2013). Although a linear fit to the data accounts for 95% of the variation in yield over this period, a two-segmented model was slightly better at describing the data (Specht *et al.*, 2014). In the two-segment model, greater gains of 29.4 kg ha^{-1} yr^{-1} were achieved after 1983 compared to gains of 21.5 kg ha^{-1} yr^{-1} before 1983 (Specht *et al.*, 2014). This faster rate from 1983 onward has been attributed to not only increased improvements in genetic and agronomic technologies, but also to greater net precipitation and higher average temperatures resulting from climate change (Twine and Kucharik, 2009).

To determine the genetic basis of historical yield improvements, cultivars with ranging year of release (YOR) have been grown in common environments (Specht *et al.*, 1999; De Bruin *et al.*, 2008; Rincker *et al.*, 2014). In the most recent experiment of this nature, Rincker *et al.*, (2014) grew 59 maturity group III cultivars with YOR dates spanning 1923 to 2007 and found linear increases of 22.8 kg ha^{-1} yr^{-1}, similar to what has been estimated in the past (Specht and Williams, 1984) and to national average U.S. yields from that same time period (Specht *et al.*, 2014). A two-segmented regression was also found to be a better fit to the yield gain, with greater rates of gain after 1964. The difference in the break points was hypothesized to be due to large confidence intervals that made it difficult to determine the actual break point (Rincker *et al.*, 2014). From these experiments, it is estimated that approximately two-thirds of past yield

gain can be attributed to genetic improvements (Specht *et al.*, 2014). While soybean yields have been increased through traditional breeding efforts, the physiological mechanisms underlying past yield gains in the U.S. are largely unknown.

Global human population is expected to increase to approximately 9 billion by 2050 (Population Division of the Department of Economic and Social Affairs of the United Nations Secretariat, 2011). If current dietary and bioenergy trends continue, this population growth is expected to require at least a doubling of grain yield globally to meet increased demand (Tilman et al., 2011). Although soybean yields have nearly quadrupled over the past century (USDA NASS), the current pace of crop improvement is not enough to reach the 2050 target of doubling crop yield (Ray *et al*., 2013). Therefore, there have been urgent calls for new strategies for increasing crop yields (Rosegrant and Agcaoili, 2010; Phillips, 2010; Godfray *et al.*, 2010). One strategy to identify promising targets for future improvement is to better understand the physiological basis of past yield improvements.

Yield potential (Y_p) is the maximum yield achieved under growth conditions devoid of abiotic and biotic stresses (Evans and Fischer 1999). Average farm yields, however, fall short of the estimated Y_p as avoidance of all environmental stresses is almost impossible to achieve and soil quality varies greatly across crop acreage (Lobell *et al.*, 2009). Y_p can be parameterized by tracking energy transfer from the sun to the seed through an equation of efficiencies adapted from Monteith (1977):

$$Y_p = 0.487 S_t \cdot \varepsilon_i \cdot \varepsilon_c \cdot \varepsilon_p.$$

In this equation, S_t is the total solar radiation that is incident upon the plant canopy throughout the growing season, 48.7% of which is within photosynthetically active spectral range. The light interception efficiency (ε_i) is the percent of that incident S_t that is intercepted by the plant canopy and is determined by the rate of canopy closure, duration of the growing season, and the stand capacity of the canopy. Energy conversion efficiency (ε_c) is the effective utilization of the intercepted radiation to produce biomass and is driven by combined gross photosynthesis minus respiratory losses of carbon throughout the canopy. Finally, the partitioning efficiency (ε_p), or harvest index, is the proportion of biomass that is allocated to reproductive versus vegetative structures (Zhu *et al.*, 2010).

The Monteith equation provides insight into the physiological mechanisms governing yield formation, and therefore, is commonly used to identify potential methods of further yield

advancement (Loomis and Amthor, 1999; Reynolds *et al.*, 2000; Parry *et al.*, 2011; Reynolds *et al*., 2011; Zhu *et al*., 2010; Ainsworth *et al*., 2012). The green revolution has led to the improvement of traits that have largely maximized ε_i and ε_p (Evans, 1993; Hay, 1995), but ε_c is below the theoretical C_3 maximum (Zhu *et al.*, 2010). Improving ε_c has been heavily focused upon recently in the search for potential strategies to enhance yield gains in the future (Amthor, 2010; Zhu et *al.*, 2010; Parry *et al.*, 2011; Raines, 2011; Ainsworth *et al.*, 2012).

Because ε_c is the combined gross photosynthesis minus respiratory losses of carbon throughout the canopy, it is hypothesized that improvements in photosynthetic carbon gain should improve ε_c, and therefore, seed yield (Zhu *et al.*, 2010). Two arguments against the notion that increased photosynthesis will improve crop yields are: (1) leaf level photosynthesis does not always correlate with yield (Moss and Musgrave, 1971; Evans, 1997; Kumudini, 2002); and (2) soybean is predominantly sink limited, which would render increased photosynthetic activity inconsequential (Borras *et al.* 2004). However, recent studies in soybean showed leaf photosynthesis correlated with yield in Chinese and Canadian germplasm (Jin *et al*. 2010; Morrison *et al*. 1999), which provides evidence against the first argument. Further, high CO_2 studies have provided evidence that traditional breeding selects a sink capacity that is greater than source capacity (Ainsworth *et al.* 2004), and increased photosynthate at elevated CO_2 leads to increased crop yields (Ainsworth and Long 2005).

Several of the suggested strategies for improving photosynthesis involve manipulating the expression of genes involved in carbon metabolism (Sinclair *et al.*, 2004; Long *et al.*, 2006; Zhu et *al.*, 2010; Parry *et al.*, 2011; Raines, 2011). Similarly, it is proposed that the alteration of "yield enhancement genes" (YEG) has the potential to increase plant growth and yield (Van Camp, 2005; Gonzalez *et al.*, 2009). Using the molecular toolbox that has been generated for *Arabidopsis thaliana*, research groups are able to perform large-scale mutant screens in search of genes that play roles in biomass accumulation or yield enhancement (Gonzalez *et al.* 2009, Sulpice *et al.* 2009). From this effort, gene targets for yield improvements in agronomically important crops have arisen, but few studies have tested these targets in crop species. Although there has been limited success of translating increases in biomass and yield in model species to field crops (Sinclair *et al.*, 2004), correlations between transcript abundance of YEG and crop yields have been reported (Yin *et al.*, 2010; Preuss *et al.*, 2012), providing further evidence that the alteration of single genes may lead to greater yields.

The aim of this thesis research was to identify mechanisms underlying past yield improvements in maturity group III soybeans by examining alterations in whole-plant physiology, leaf-level carbon metabolism, and transcript abundance within historic cultivars. The objectives of the first study were to quantify how the Monteith efficiencies have been altered by traditional breeding and to determine the current status of each parameter in relation to its proposed theoretical maxima. The second study investigated changes in photosynthetic and respiratory capacity in historic cultivars, and the third study investigated how the expression of YEG changed with cultivar year of release. Together these chapters provide understanding of how soybean breeding over the past 80 years has altered soybean physiology and identify potential targets for improving future soybean seed yield.

CHAPTER II: HISTORICAL GAINS IN SOYBEAN (*GLYCINE MAX* MERR.) SEED YIELD ARE DRIVEN BY LINEAR INCREASES IN LIGHT INTERCEPTION, ENERGY CONVERSION, AND PARTITIONING EFFICIENCIES[1]

Introduction

Soybean (*Glycine max*) yields have steadily increased throughout the past century from advances made in breeding, improved management practices, and increased atmospheric carbon dioxide concentrations (Specht *et al*., 1999; De Bruin *et al.*, 2008; Rowntree *et al.*, 2013). However, the current rate of gain is insufficient to meet the United Nations target of doubling crop yields by 2050 in order to meet the needs of a growing population (Tilman *et al.*, 2011; Ray *et al.*, 2013). While soybean yields have been increased through traditional breeding efforts, the physiological mechanisms underlying past yield gains in the U.S. are largely unknown. An understanding of the physiological basis of past improvements in soybean yield could help identify strategies for increasing future production.

Yield potential (Y_p) is defined as the maximum yield achieved when a crop is grown in absence of biotic and abiotic stresses (Evans and Fischer, 1999). Y_p can be parameterized by different efficiencies in the following equation adapted from Monteith (1977):

$$Y_p = 0.487 S_t \cdot \varepsilon_i \cdot \varepsilon_c \cdot \varepsilon_p.$$

In this equation, S_t is total incident solar radiation during the growing season of which ~48.7% is photosynthetically active. Light interception efficiency (ε_i) is determined by the speed and duration of canopy closure along with canopy size and architecture. Energy conversion efficiency (ε_c), or radiation use efficiency, is determined by the amount of solar energy that is transformed into biomass through the balance of photosynthesis and respiration. Partitioning efficiency (ε_p), or harvest index, is determined by the amount of biomass energy allocated to vegetative versus reproductive structures (Zhu *et al*., 2010). The Monteith equation tracks energy

[1] This chapter appeared in its entirety in the Journal of Experimental Botany and is referred to later in this dissertation as "Koester *et al.*, 2014". **Koester RP, Skoneczka JA, Cary TR, Diers BW, Ainsworth EA.** 2014. Historical gains in soybean (*Gylcine max* Merr.) seed yield are driven by linear increases in light interception, energy conversion, and partitioning efficiencies. *Journal of Experimental Botany* **65**, 3311-3321**.** This article is reprinted with the permission of the publisher and is available online at http://www.oxfordjournals.org/en/ using doi: 10.1093/jxb/eru187.

transfer from the sun to the seed and provides insight into the physiological mechanisms that ultimately govern yield potential. As a result, the Monteith equation has been used to assess which parameters are at their theoretical maxima and which could be improved further to advance yield (Gifford *et al*., 1984; Loomis and Amthor, 1999; Reynolds *et al.*, 2000; Reynolds *et al*., 2010; Zhu *et al*., 2010; Ainsworth *et al*., 2012).

The extent to which soybean breeding strategies have improved ε_i, ε_c, and ε_p in U.S. soybean germplasm has not been investigated. In Chinese and Canadian soybean germplasm, negative correlations between plant height and lodging score with cultivar year of release (YOR) have been reported (Jin *et al*., 2010; Morrison *et al*., 2000). These changes in height and lodging improved the standing power of the crop and are hypothesized to increase ε_i (Zhu *et al*., 2010). Improved ε_p with YOR in Chinese and Canadian germplasm was attributed to increased seed biomass with little or no increase in total above-ground biomass (Jin *et al*., 2010; Morrison *et al*., 1999). There is some evidence that ε_c also has been improved by breeding because leaf-level photosynthetic carbon assimilation increased with YOR (Jin *et al.*, 2010; Morrison *et al.*, 1999). However, ε_c is the season-long balance between C gain and C loss, and changes in carbon utilization and respiration can offset changes in photosynthesis. Additionally, a direct correlation between leaf-level photosynthesis and crop yield is not consistently apparent (Kumudini, 2002). Therefore, it is not known how decades of soybean breeding have altered ε_c.

It has been suggested that modern cultivars in high-yielding environments achieve theoretical maximum efficiencies of ε_i (0.9) and ε_p (0.6), while ε_c is below the theoretical C_3 maximum (0.094; Zhu *et al*., 2010). However, there has not been a comprehensive study that parameterizes the Monteith equation across U.S. soybean cultivars with a range of release dates in order to assess how decades of breeding have altered the efficiencies in the field. Further, there is insufficient knowledge about whether elite germplasm are reaching their theoretical maximum efficiencies. Therefore, in order to elucidate the physiological mechanisms of yield improvement in historical soybean germplasm this study parameterizes the Monteith equation in U.S. soybean cultivars released from 1923-2007. It is hypothesized that: (1) breeding has increased canopy duration and decreased lodging, therefore ε_i will increase with cultivar YOR; (2) breeding has improved net C balance in soybean, therefore ε_c will increase with cultivar YOR; (3) seed yield has been increased by traditional breeding while vegetative biomass has not been affected, therefore ε_p will increase with YOR.

Materials and methods

Experimental design

Research was conducted at the Crop Research and Education Center in Urbana, IL (40°N, 88°14`W) in 2012 and 2013. Twenty-four indeterminate, maturity group III soybean cultivars were chosen to represent 84 years of past yield gains (Table 1). The publically developed cultivars were obtained from the USDA Soybean Germplasm Collection, Urbana, IL, courtesy of Dr. Randall Nelson. Non-public selections were obtained from Pioneer Hi-Bred, Syngenta, and Monsanto and were coded as private entries. Cultivars were chosen to minimize differences in maturity date and to maximize evenness of distribution across the years of cultivar release. Seed of all cultivars were produced in a common environment in Illinois the year prior to each experiment. Each year of the experiment was arranged in a randomized complete block design with three replicates. In one block, the cultivars were each grown in large plots (3.05 m x 12.20 m with 16 rows in 2012 and 3.05 m x 9.44 m with 12 rows in 2013) and in the two remaining blocks, cultivars were grown in smaller plots (3.05 m x 3.05 m with 4 rows in both years). The smaller plots were used to determine seed yield at maturity as well as lodging while the larger plots were used for destructive physiological measurements, tissue sampling, as well as yield determination at maturity. Experimental plots were planted at a row width of 0.76 m and thinned after emergence to a uniform density (Table 2) after unequal stand density was observed in 2011 in a preliminary experiment (Fig. 2.1A). Unequal stand density was caused by differences in germination rates (Fig. 2.1B). Daily meteorological data including S_t (Fig. 2.2A, B), temperature (Fig. 2.2C, D), and precipitation (Fig. 2.2E, F) were collected ~1.5 km from the field site by the Illinois Climate Network monitoring station (Fig.2.2; Angel, 2009). Plots were irrigated using drip-line tubing four times during the 2012 season to prevent water stress (Fig 2.2E). Drip-line tubing was not laid in 2013 because of ample precipitation early in the growing season.

Light interception and conversion efficiency

Measurements of ε_i were made once or twice per week throughout the growing season. The photosynthetically active radiation (PAR) was measured above (I_a) and below (I_b) the canopy in two undisturbed areas in each large plot between 11:00 and 14:00 on clear-sky days with a 0.87 m line quantum sensor (AccuPAR LP-80, Decagon Devices, Pullman, WA, USA). ε_i was

estimated from two measurements of PAR directly above the canopy and eight measurements below the canopy. Below-canopy measurements were made ~2.5 cm above the ground across a 0.76 m transect between rows. ε_i was then calculated as $1 - \frac{I_a}{I_b}$ (Nobel *et al.*, 1993). The season-long average ε_i was calculated by averaging all measurements taken throughout the season. ε_i measurements were stopped and assumed to be 0 once the plot reached growth stage R7 defined by pod maturity (Fehr *et al*., 1971), by which time most of the remaining foliage had senesced.

Above-ground biomass accumulation per unit area was measured every two weeks. Avoiding the edges of the plot (0.5 m), a 1 m length of row was harvested at 2.5 cm above the ground. Plants were counted and separated into leaf, stem (including petioles and petiolules), and pod sections. Plant material was then dried for one week at 70°C and weighed. In order to convert total biomass into energy equivalents, seeds, leaves, and stems were ground and analyzed for total energy content using adiabatic bomb calorimetry (model 1261, Parr Instrument Co., Moline, IL, USA) with benzoic acid as a standard (Figs. 2.3, 2.4). Because biomass measurements were made in parallel with ε_i measurements, cumulative intercepted radiation was calculated as the accumulated intercepted PAR (PAR_i) multiplied by the average ε_i for each period between point estimates of ε_i leading up to each biomass harvest. For calculation of season-long ε_c, cumulative PAR_i (MJ m^{-2}) was plotted against cumulative biomass energy (MJ m^{-2}) until peak biomass was observed. The slope of the linear fit was used to estimate ε_c (Monteith, 1972) and it was assumed that ε_i was 0 on the day of crop emergence.

Partitioning efficiency and yield

ε_p was calculated as the ratio of seed mass to total above-ground biomass at full maturity (R8; Fehr *et al*., 1971). Total seed and stem biomass was measured as previously described above, except 2 m of row were harvested for calculation of ε_p. Lodging scores were determined in all three experimental plots using a 0-10 scale according to the following system: most main stems were completing vertical at 0° (0), 45° (5), completely horizontal at 90° (10). When the cultivars had reached maturity, yield was determined by harvesting two center rows from each of the three yield plots with a 2-row combine and estimates were adjusted to 13% seed moisture content.

Statistical analysis

A significant correlation between yield, Monteith efficiencies, and cultivar YOR was tested using least squares regressions (PROC MIXED procedure, SAS version 9.2, SAS Institute Inc., Cary, NC, USA) or first order linear regression (SigmaPlot, Systat Software, Inc, Richmond, CA, USA). A t-test was used to determine if linear regressions slopes were significantly different among years. A two-segment linear regression model (PROC NLMIXED procedure, SAS version 9.2, SAS Institute Inc., Cary, NC, USA) was also fit to the data and compared to the linear fit using the Akaike information criterion coefficient.

Results

Yield increased linearly with cultivar YOR

There was a linear improvement in soybean yields with cultivar YOR, with increases of 32.1 kg ha^{-1} yr^{-1} in 2012 and 20.8 kg ha^{-1} yr^{-1} in 2013 (Fig. 2.5A, B). The rate of yield gain in 2012 was significantly greater than in 2013 ($p < 0.005$). Older cultivars showed less year to year variation in seed production, with insignificant yield differences between years, while the newest cultivars significantly yielded more in 2012 compared to 2013 (Fig. 2.5A, B). Newer cultivars (Private 3-14, Private 3-13 and IA3023) were consistently among the highest yielding and older cultivars (Dunfield and Illini) were the lowest in both years of the experiment.

ε_i increased with cultivar YOR

Season-long average ε_i increased with YOR in both 2012 and 2013 (Fig. 2.5C, D), and the slopes in the trends were not significantly different between years ($p = 0.24$). Increases in season-long ε_i with cultivar YOR were driven by a longer growing season with more recent cultivars maturing later (Fig. 2.6). The growing season was longer in lines released in the 1980's – 2000's, compared to the lines released in the 1920's – 1940's (Fig. 2.6). There wasn't a difference in the rate of canopy closure in older or newer cultivars, and most cultivars approached 90% closure by ~60 d after planting (Fig. 2.6). Lodging, which negatively affects ε_i at the end of the growing season, also decreased with YOR (Fig. 2.7).

ε_c increased with cultivar YOR

ε_c increased with cultivar YOR in 2012 and 2013 (Fig. 2.5E, F). Improved ε_c was driven by greater biomass accumulation per PAR_i in more recently released lines (Fig. 2.8). In 2012,

cultivars released between 1990 and 2000 accumulated 14.1 MJ m^{-2} in biomass over the growing season, compared to 12.9 MJ m^{-2} in cultivars released between 1920 and 1940. Similarly, in 2013, cultivars released between 1990 and 2000 accumulated 17% more biomass over the growing season compared to cultivars released between 1920 and 1940.

While the slopes of the trends in ε_c with cultivar YOR did not significantly differ between years (Fig. 2.5E, F), average values of ε_c were significantly greater in 2013 compared to 2012 (Fig. 2.5E, F). This was driven primarily by differences in accumulated PAR in the two years. In 2012, cultivars accumulated ~13% more total peak biomass than in 2013, but did so by using 33% more intercepted PAR resulting in lower values of ε_c (Fig. 2.8).

ε_p increased with cultivar YOR

ε_p increased significantly with cultivar YOR in both years of the study (Fig. 2.5G, H). Gains in ε_p were driven primarily by increases in total seed biomass as ~80% of the gain in total biomass was caused by increases in seed biomass (Fig. 2.9). Although the values of seed and total biomass were greater in 2012 compared to 2013 (Fig. 2.9), the ratio of seed to total biomass was similar, and therefore, the rate of gain in ε_p with YOR was the same in both years (Fig. 2.5G, H).

Yield correlations with Monteith efficiencies

In 2012, all three Monteith efficiencies (ε_i, ε_c, and ε_p) significantly correlated with yield (Fig. 2.10), and ε_c and ε_p were correlated with one another (Fig. 2.10). However, ε_i was not correlated with ε_c or ε_p in 2012. In 2013, ε_i and ε_p significantly correlated with yield (Fig. 2.11) but ε_c did not (Fig. 2.11). ε_i was more strongly correlated to yield in 2013 (Fig. 2.11), a year with ~30% less total solar radiation over the growing season compared to 2012, and ε_c was not significantly correlated with yield in 2013 (Fig. 2.11). ε_p is auto-correlated with seed yield, and therefore showed very high correlation coefficients in both years (Figs. 2.10, 2.11).

Discussion

In field trials of U.S. soybean cultivars released over the past 84 years, seed yield significantly increased with YOR. When separating yield into its physiological efficiencies, there were consistent increases in the efficiencies by which canopies intercepted solar energy, converted it

into biomass, and partitioned biomass into yield. In a highly productive agricultural area in the Midwest U.S., peak ε_i is >90% and ε_p is reaching the theoretical maxima value (60%) in recently released soybean cultivars. However, there is still room for further improvements in ε_c in modern soybean cultivars.

This study of historical soybean cultivars estimated rates of soybean yield gain of 32.1 kg ha^{-1} yr^{-1} in 2012 and 20.8 kg ha^{-1} yr^{-1} in 2013. These rates are in line with the gains reported in a literature review by Specht *et al*. (1999) and are similar to rates reported in a recent study of 60 cultivars with a similar range of YOR dates that also included the 24 cultivars grown in this present study (Rowntree *et al*., 2013; Rincker *et al.*, 2014). While Rincker *et al.* (2014) found the data were better described by a two-segment linear fit with different slopes before and after 1964, the rates of yield gain in this study were better described by a single linear fit, perhaps because there was less power in this study to detect differences in the rate of yield gain before and after 1964. The gains in soybean yield reported here are also similar to improvements reported for other major crops including maize (*Zea mays*; Duvik and Cassman, 1999; Richards, 2000), rice (*Oryza sativa*; Peng *et al*., 2000) and wheat (*Triticum aestivum*; Reynolds *et al*., 1999; Shearman *et al*., 2005). The greater rates of yield gain observed in 2012 compared to 2013 were likely caused by differences in environmental factors and irrigation. The experimental site experienced hot, dry growing conditions in 2012, so plots were irrigated to reduce water stress. The 2013 growing season had lower maximum temperatures, less incoming solar radiation, and ample water early in the season. However, drought conditions occurred after the canopy closed and when seeds were filling which likely contributed to the lower rate of gain in seed yield in 2013. When comparing the two years of the study, it was also notable that older cultivars had more consistent yields in 2012 and 2013, while more recently released cultivars had large variation in the two years, perhaps meaning that newer lines have greater environmental sensitivity. These results are consistent with Rincker *et al.* (2014), who found greater rates of soybean yield gain in high yielding environments, and lower yield stability in more recently released cultivars.

The effective capture of solar radiation across the growing season determines how much solar energy is available for conversion into biomass and therefore yield. In this study, ε_i increased with cultivar YOR similarly across both years with ranges of season-long average ε_i intercepting approximately 50-75% of the growing season's PAR. Peak ε_i in all soybean

cultivars was >90%, consistent with previous reports (Dermody *et al*., 2008). However, the seasonal ε_i measured in this study is lower than the theoretical maximum ε_i for soybean of ~90% (Zhu *et al*., 2010) and lower than previously reported levels of 89% (Dermody *et al*., 2008). This may be because the current study used wider row spacing than Dermody *et al.* (2008), and because the current study took more measurements early in the growing season when the canopy was still developing. There was no difference in time to canopy closure among new and old soybean varieties, but rather an increase in the duration of a photosynthetically active canopy allowing greater capture of S_t. This was in part because more recent cultivars have later maturity dates (Rowntree *et al.*, 2013), but also because lodging significantly decreased with YOR which lengthened the duration of an active canopy. Other studies in soybean have reported similar improvements in lodging score over years of breeding (Specht *et al.*, 1999; Morrison *et al.*, 2000; Jin *et al.*, 2010). There are very few direct estimates of ε_i in soybean, but leaf area index (LAI) is commonly measured and used to indicate ε_i. A decreasing trend in LAI with YOR has been reported (Morrison *et al.*, 1999; Jin *et al.*, 2010) indicating that newer cultivars with lower LAI may have reduced capacity for intercepting light. However, while LAI can be a good indicator of light interception at the early stages of canopy closure, at an LAI of 3.5-4.0, light interception exceeds 95% (Board and Harville, 1992). Therefore, LAI values above ~4.0 reveal very little about ε_i. Improvement strategies for light interception in major crops tend to focus primarily on extending the growing season and/or engineering for optimal crop canopy architecture (Reynolds *et al.*, 2000; Parry *et al.*, 2010; Zhu *et al*., 2010), which would increase the total S_t for the crop to intercept. In rice for example, each day added to the growing season translated into a 180 kg ha^{-1} increase in yield (Akita, 1988).

Energy conversion efficiency and its improvement has been the focus recently of many yield improvement strategies (Amthor, 2010; Zhu et *al.*, 2010; Parry *et al.*, 2011; Raines, 2011; Ainsworth *et al.*, 2012). Yet the extent of how ε_c has been improved through historical breeding is not well understood. In this study, ε_c increased with YOR in both 2012 and 2013 at very similar rates leading to a ~36% improvement over the 84 years covered in this study. A similar increase in ε_c in wheat cultivars released from the 1970's to 1990's has been reported (Shearman et al. 2005); however, earlier studies of different wheat cultivars failed to report a similar trend (Slafer *et al.*, 1990; Calderini *et al.*, 1997). In the current study, the average ε_c was 29% higher in 2013 compared to 2012, with a maximum ε_c of 2.9% in 2012 and 4.3% in 2013. These rates

are higher than the maximum rates of field grown C_3 crops (4.8%) reported in Zhu *et al.*, (2008), but still well below the theoretical maximum of 9.4% (Zhu *et al.*, 2010). ε_c is estimated from the linear relationship between biomass accumulation and intercepted light, and gains in ε_c in recently released soybeans came from increased biomass production for a given amount of intercepted light (Fig. 2.8). Changes in respiration or photosynthesis could underpin this trend in ε_c, and previous work in Canadian and Chinese germplasm suggests that leaf-level photosynthesis has improved with YOR (Jin *et al.*, 2010; Morrison *et al.*, 1999). However, future studies are needed to determine the mechanisms driving improvements in ε_c in these MG III historical lines. The average ε_c in 2012 was lower than in 2013, because although the crop intercepted 33% more PAR in 2012 than in 2013, peak biomass was only 13% greater in 2012 than 2013. C3 leaf photosynthesis saturates at ~30% full sunlight, and plants are not able to utilize all the intercepted solar radiation, which results in decreased efficiencies of energy conversion (Sinclair and Muchow, 1999; Ort, 2001). A recent meta-analysis by Slattery *et al.* (2013) found that in shading experiments, ε_c increased on average by 18% when plants were grown in shaded conditions compared to full sunlight. Consistent with the meta-analysis, average ε_c of soybean was greater in a year with less solar radiation; however, despite the increased efficiency in 2013, 2012 resulted in higher absolute seed yields. Although the plants were less efficient in the amount of C fixed per MJ of light in 2012, the plants had higher rates of incident solar radiation throughout the growing season which more than compensated for the loss of efficiency.

Consistent increases in ε_p with YOR were observed in 2012 and 2013.The range of ε_p for both years was similar (~0.3-0.55), and the most recently-released cultivars have values approaching the proposed theoretical maximum of 0.60 (Hay, 1995; Zhu *et al.*, 2010), although the data here do not provide evidence that ε_p in soybean is reaching a plateau. The linear improvement of ε_p with YOR was achieved through tripling seed biomass per area but only doubling total biomass per area (Fig. 2.9). The rate of gain in ε_p in Chinese soybean germplasm was similar (Jin *et al.*, 2010). In Canadian soybean germplasm, historical improvements in ε_p were only due to increases in seed weight and not total biomass (Morrison *et al.*, 1999). In other major food crops, particularly small grains, improvements in ε_p largely drove improvements in yield from 1900 to 1980 (Hay, 1995). In wheat, linear increases in ε_p were found with YOR in

the UK and in Mexico, and were achieved through increased grain yield with no increase in total biomass (Austin *et al.*, 1989; Sayre *et al.*, 1997). More recently, Shearman *et al*. (2005) reported that ε_p leveled off at ~0.5 when they looked at cultivars of wheat that spanned the 1970-1995. Historically, rice showed improvements in ε_p until it reached a maximum of around 0.6 in the 1980's when increases in yield were then attributed to greater rates of biomass production (Hay, 1995; Peng *et al.*, 2000). The ε_p of maize was already high (~0.45) in the early 1930s, and therefore gains in maize yield were made through increases in total biomass (Hay, 1995; Richards, 2000).

The contribution to yield gain by the Monteith efficiencies was investigated by analyzing their correlations with yield. ε_p is auto-correlated with yield and not surprisingly showed the strongest correlations in both 2012 and 2013 (Figs. 2.10, 2.11). Yield correlations with ε_i and ε_c were more variable and weaker. Interestingly there was no correlation between ε_i and ε_c, suggesting that the improvements in these traits in historical germplasm may have been independent. The correlations with yield suggest that improvements in all Monteith efficiencies were important to past yield gains, and they are all targets of international efforts to improve future C_3 crop yields (Reynolds *et al*., 2010).

Conclusion

Several physiological changes have accompanied the impressive gains in soybean yield over the past 80 years. First, soybean canopies of more recently released cultivars have greater season-long canopy interception efficiencies owing to longer growing seasons and improved resistance to lodging. Second, modern soybean cultivars have better efficiencies of converting light energy into above-ground biomass and produce 9-17% more above-ground biomass energy in a growing season than cultivars released before 1950. Third, the partitioning of carbon to seeds has greatly improved in modern soybean lines.

Tables and Figures

Table 2.1. List of maturity group III soybean cultivars grown with respective year of release (YOR) dates and plant introduction (PI) number.

Cultivar	YOR	PI No.
Dunfield	1923	PI548318
Illini	1927	PI548348
AK (Harrow)	1928	PI548298
Mandell	1934	PI548381
Lincoln	1943	PI548362
Adams	1948	PI548502
Ford	1958	PI548562
Shelby	1958	PI548574
Ross	1960	PI548612
Adelphia	1964	PI548503
Wayne	1964	PI548628
Calland	1968	PI548527
Williams	1971	PI548631
Woodworth	1974	PI548632
Zane	1984	PI548634
Private 3- 2	1986	n/a[†]
Resnik	1987	PI534645
Private 3- 9	1989	n/a
Private 3-19	1994	n/a
Private 3-11	1996	n/a
IA 3010	1998	n/a
IA 3023	2003	n/a
Private 3-13	2004	n/a
Private 3-14	2007	n/a

[†]not available

Table 2.2. Summary of meteorological conditions, plant density and planting and harvest dates in the two years of study.

Year	Planting Date	Harvest Date	Final plant Density (plants ha^{-1})	Precipitation (mm)	Average Max Temperature (°C)	Radiation (MJ m^{-2})
2012	12 May	30 Oct	386,421	483[†]	30.6	2,944
2013	16 May	14 Oct	379,325	315	28.1	2,130

[†]Precipitation + irrigation

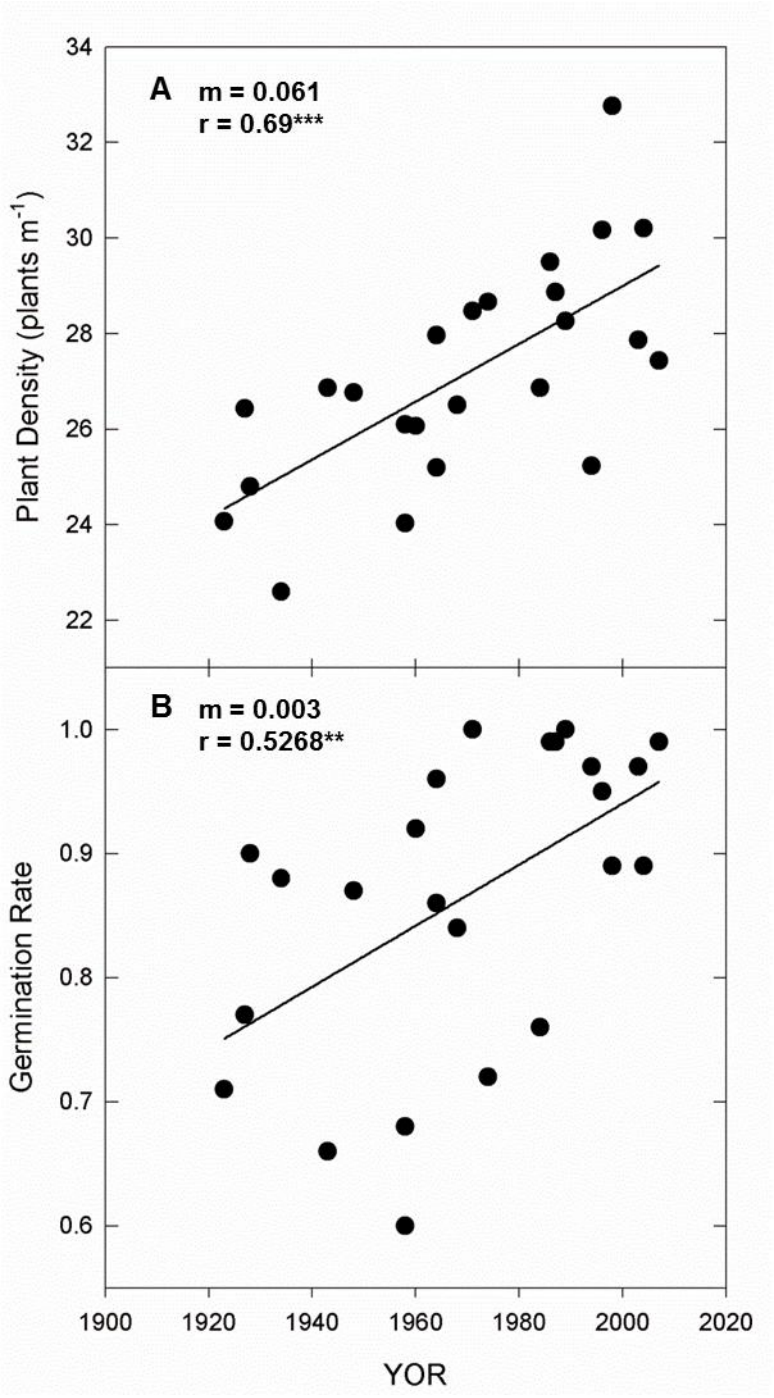

Figure 2.1. Plant density and seed germination versus YOR in 2011. Planting density (A) and germination rate (B) plotted against cultivar YOR. All lines represent the least squares regression (**p < 0.01, *** p < 0.001).

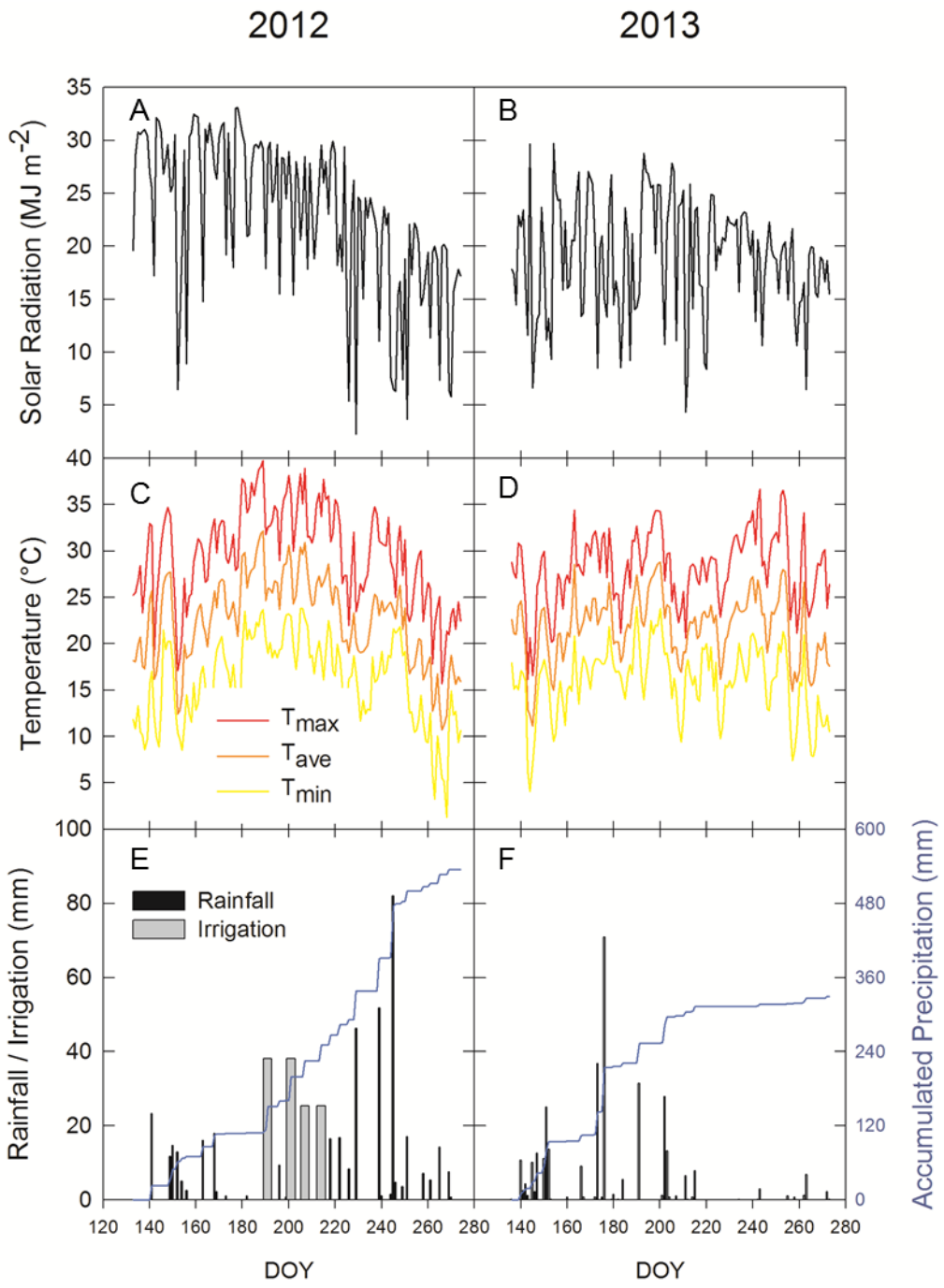

Figure 2.2. Meteorological data for the 2012 and 2013 experimental growing seasons. Meteorological conditions collected for the 2012 and 2013 growing seasons (planting date to 30-Sep). Panels (A-B) show the daily total solar radiation. Also shown are the daily maximum (red), average (orange), and minimum (yellow) temperatures (C-D). Rainfall (black bars) and irrigation (grey bars) events are shown (E-F) along with the accumulated precipitation across the growing season (blue, E-F).

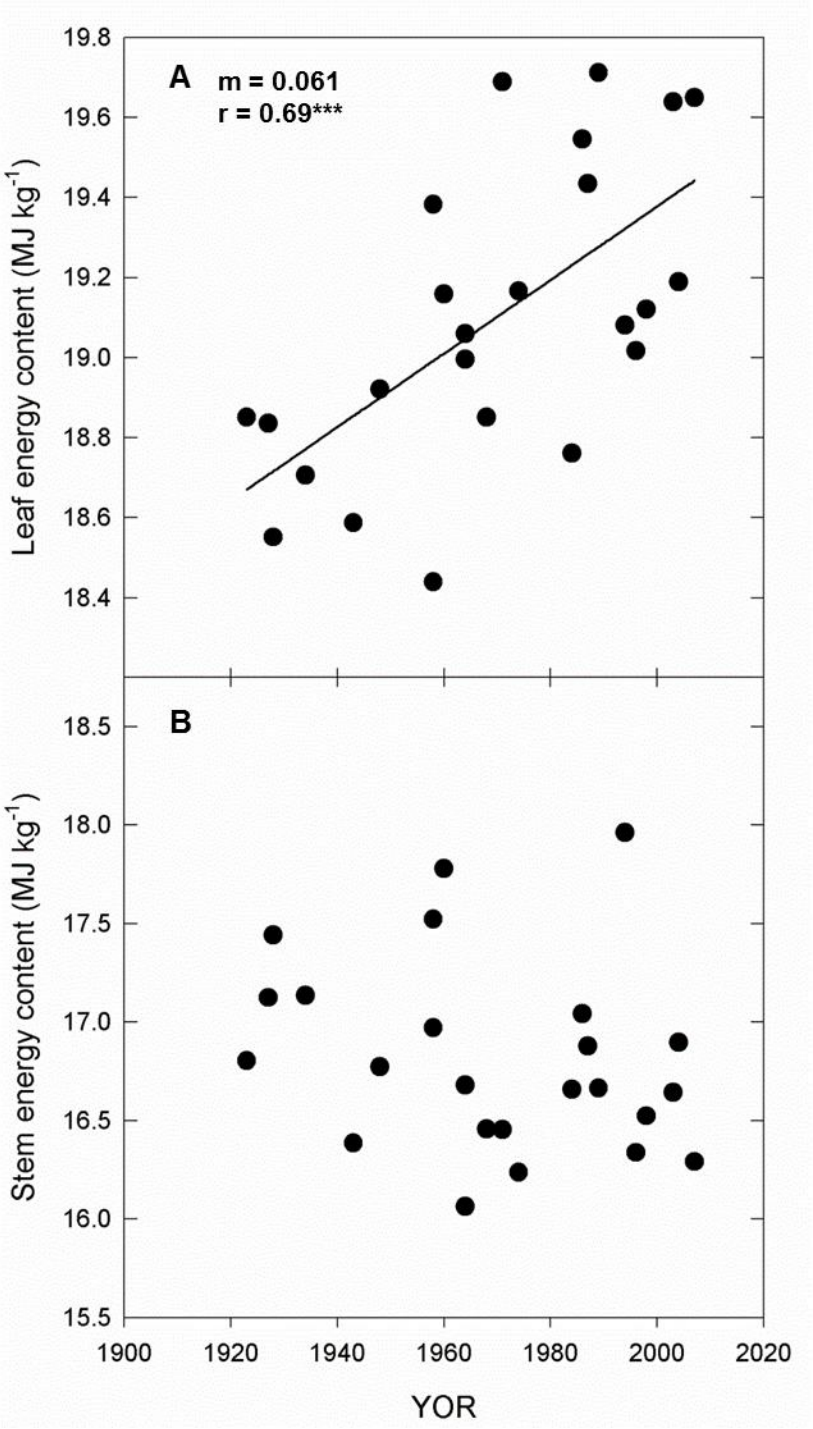

Figure 2.3. Leaf and stem energy content versus YOR. Leaf and stem energy content is plotted against cultivar YOR with the line representing the least squares regression (*** $p < 0.001$).

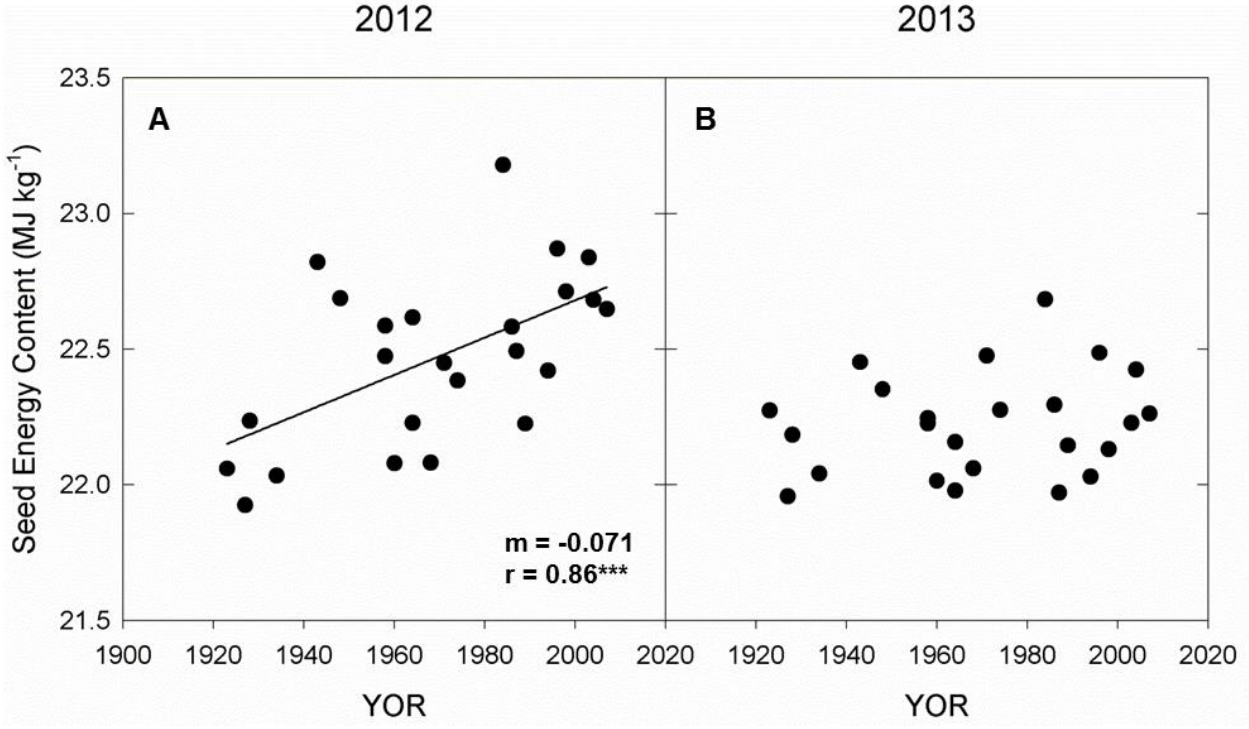

Figure 2.4. Seed composition versus YOR in 2012 and 2013. The energy content (A, B) of the seed is shown plotted against YOR in 2012 and 2013. All lines represent the least squares regression (*** $p < 0.001$).

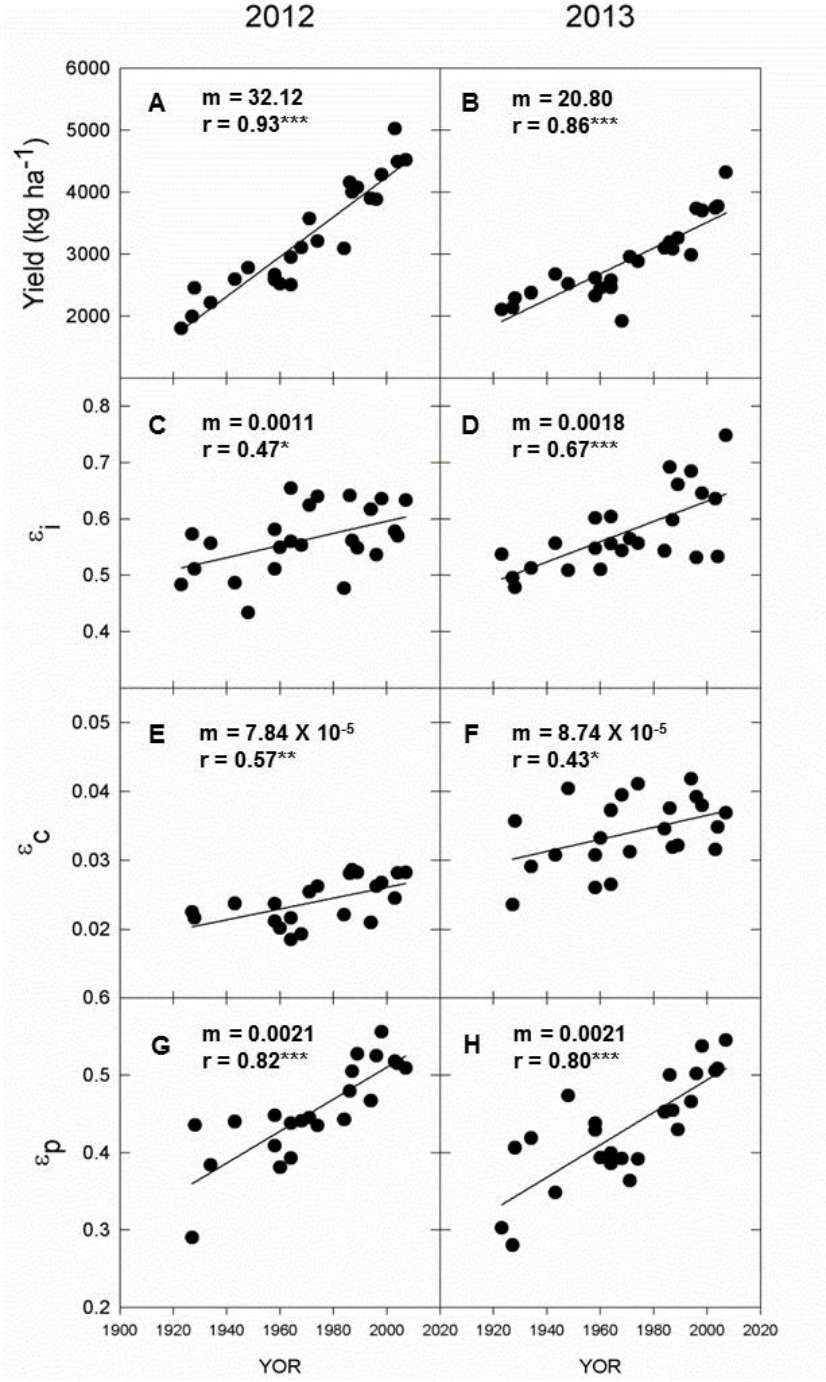

Figure 2.5. Historical gains in seed yield, ε_i, ε_c and ε_p with soybean cultivar year of release (YOR). Seed yield (A-B), seasonal average interception efficiency (ε_i, C-D), conversion efficiency (ε_c, E-F) and partitioning efficiency (ε_p, G-H) are shown plotted against cultivar YOR for the 2012 and 2013 growing seasons. Lines represent the significant least squares regression. The slope (m) and the Pearson correlation coefficient (r) are reported and *, ** and *** denotes significance at $p < 0.05$, $p < 0.01$ and $p < 0.001$).

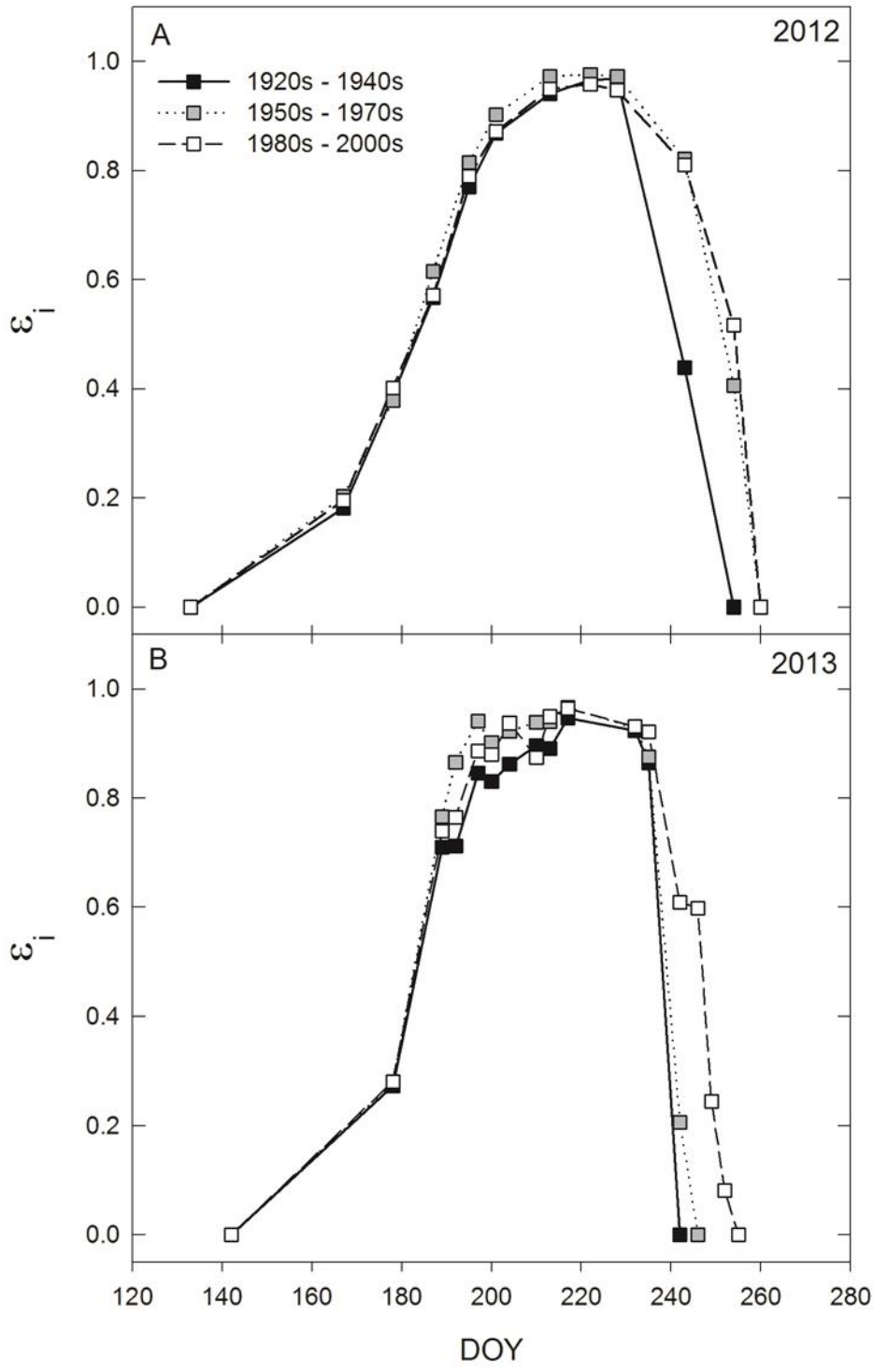

Figure 2.6. Interception efficiency (ε_i) across the growing season in 2012 and 2013. Point measurements of ε_i are plotted across the 2012 (A) and 2013 (B) growing season for each of the 24 soybean cultivars. Soybean cultivars are grouped by YOR within the following time periods: 1920s-1940s (black squares), 1950s-1970s (gray squares) and the 1980s-2000s (white squares).

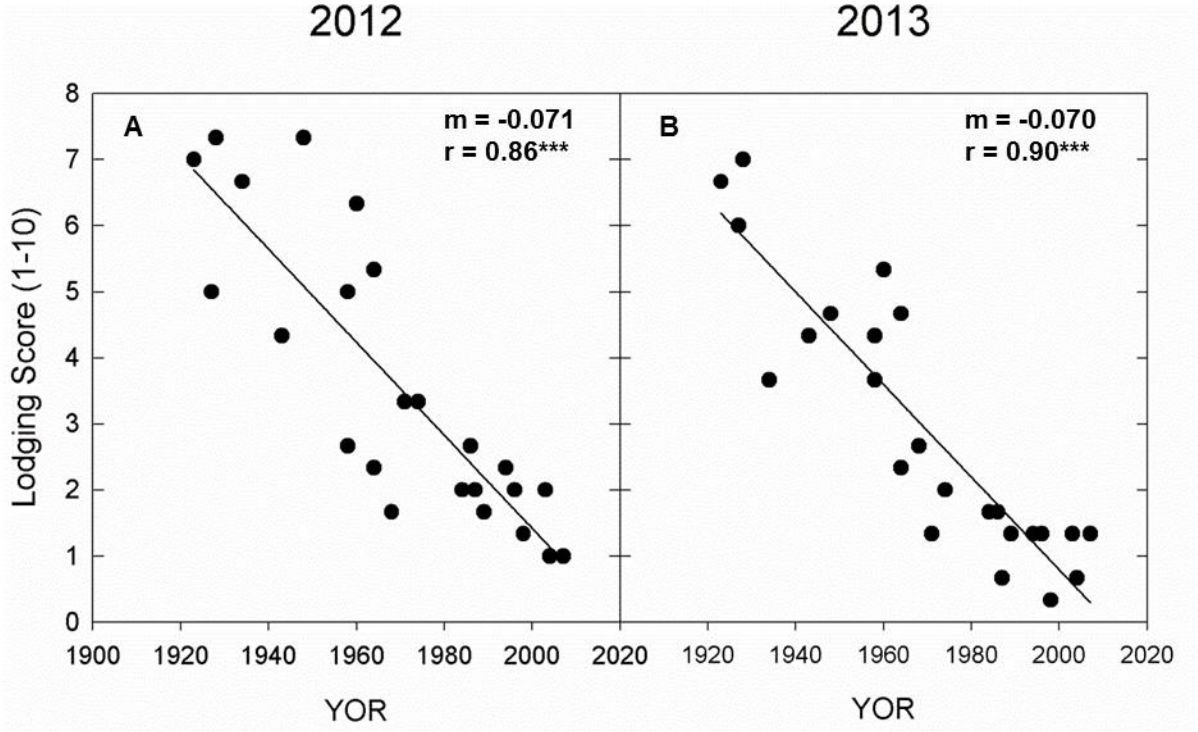

Figure 2.7. Lodging score versus YOR across the 2012 and 2013 growing seasons. Lodging score is plotted against cultivar YOR with the line representing the least squares regression (*** $p < 0.001$). Each point is the average of three replicates.

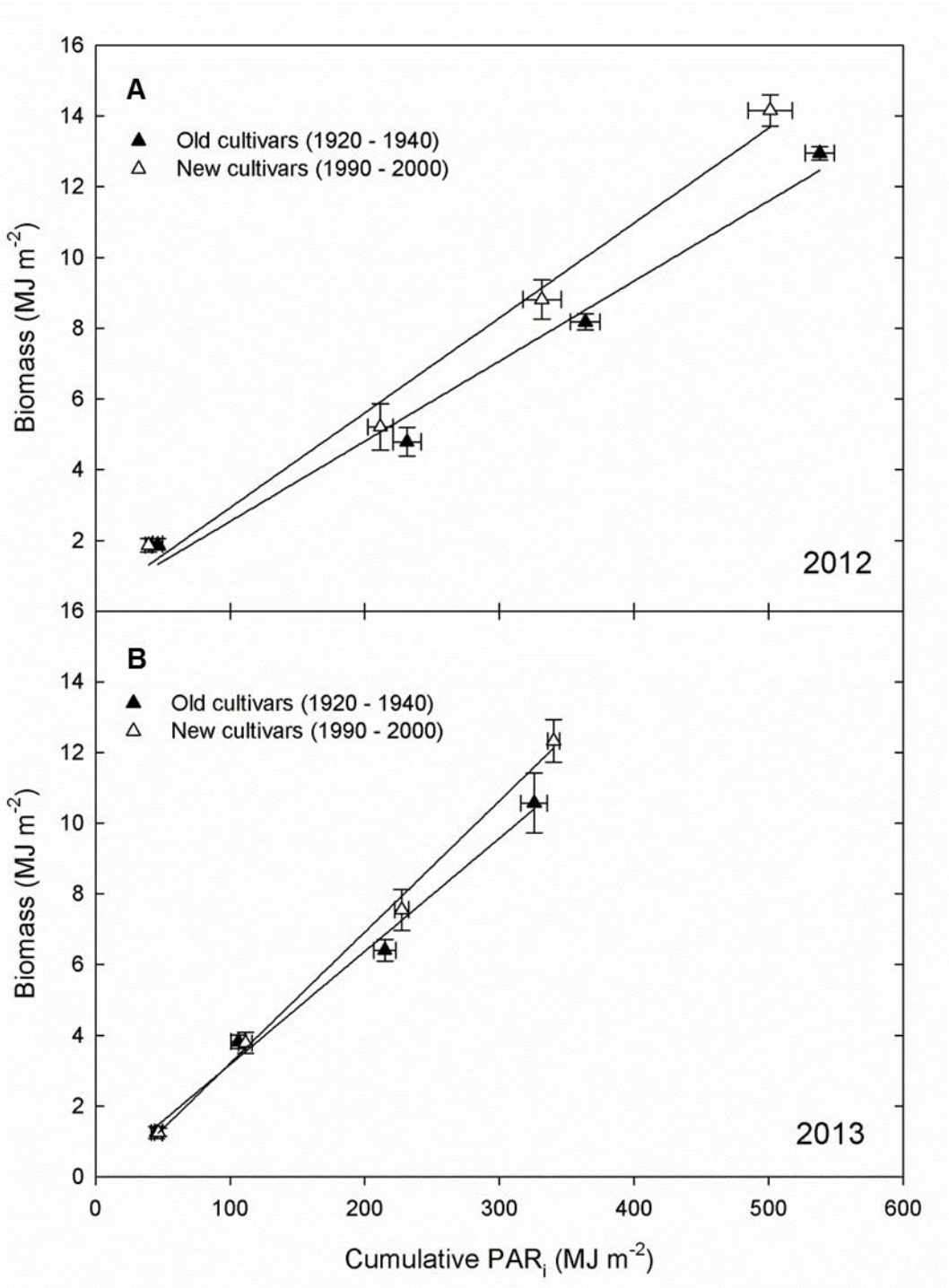

Figure 2.8. Accumulated above-ground biomass versus cumulative PAR_i. Lines shown are the least squared regression between dry biomass versus cumulative PAR_i. The slope of the line is ε_c. Each point represents the average biomass and cumulative PAR_i for the five oldest cultivars (triangles) and the five most recently released cultivars (open triangles) in 2012 (A) and 2013 (B).

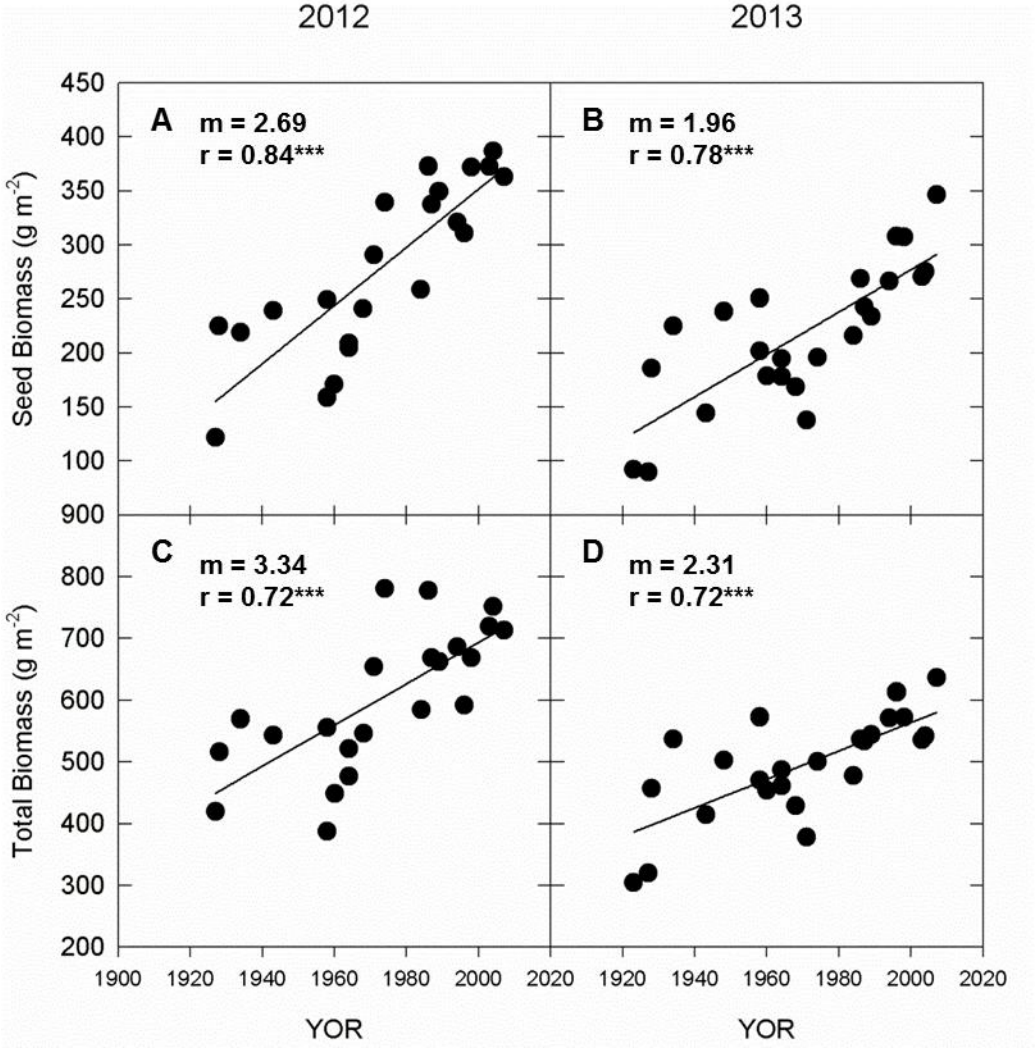

Figure 2.9. Determinants of partitioning efficiency (ε_p) versus YOR. Seed biomass (A-B) and total biomass (C-D) at growth stage R8 plotted against cultivar YOR in 2012 and 2013. Lines represent the significant least squares regression (*** $p < 0.001$).

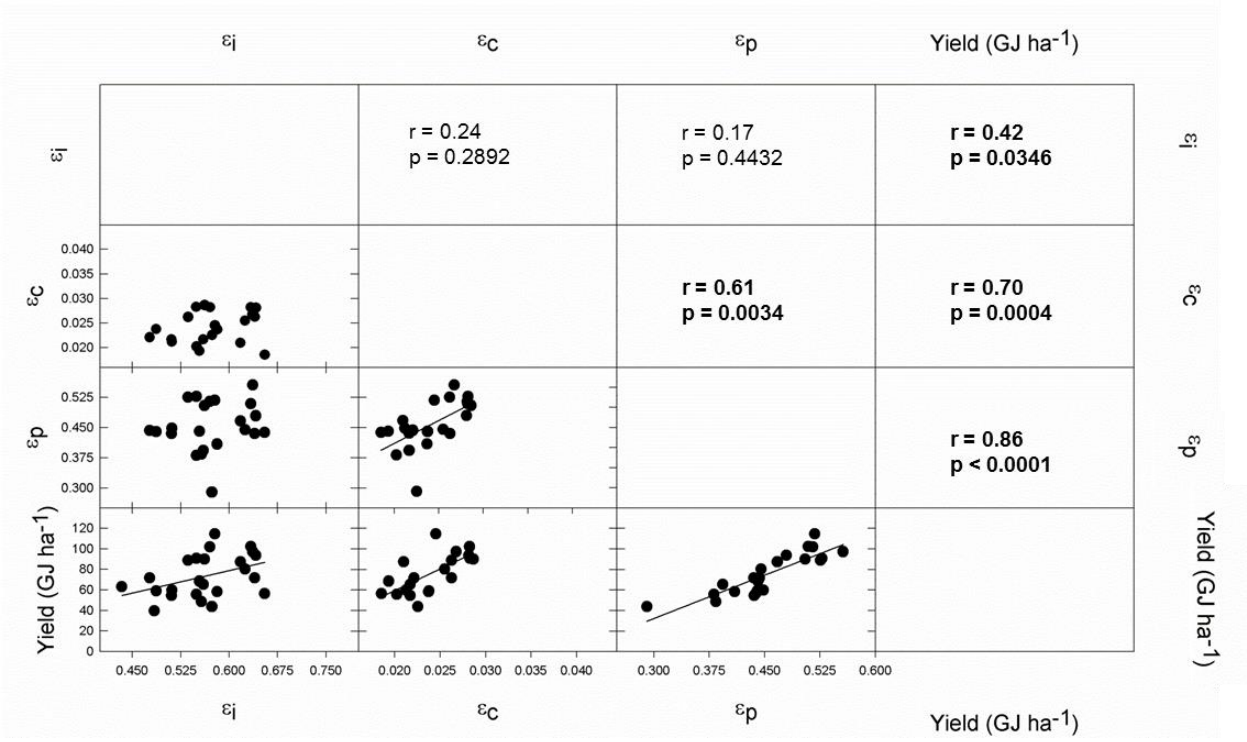

Figure 2.10. 2012 correlation matrix of yield and the Monteith efficiencies. Scatterplots and correlation coefficients are plotted in a matrix where lines represent the significant least squares regression. The Pearson correlation coefficients (r) and p-values (p) are reported and shown in bold when significant.

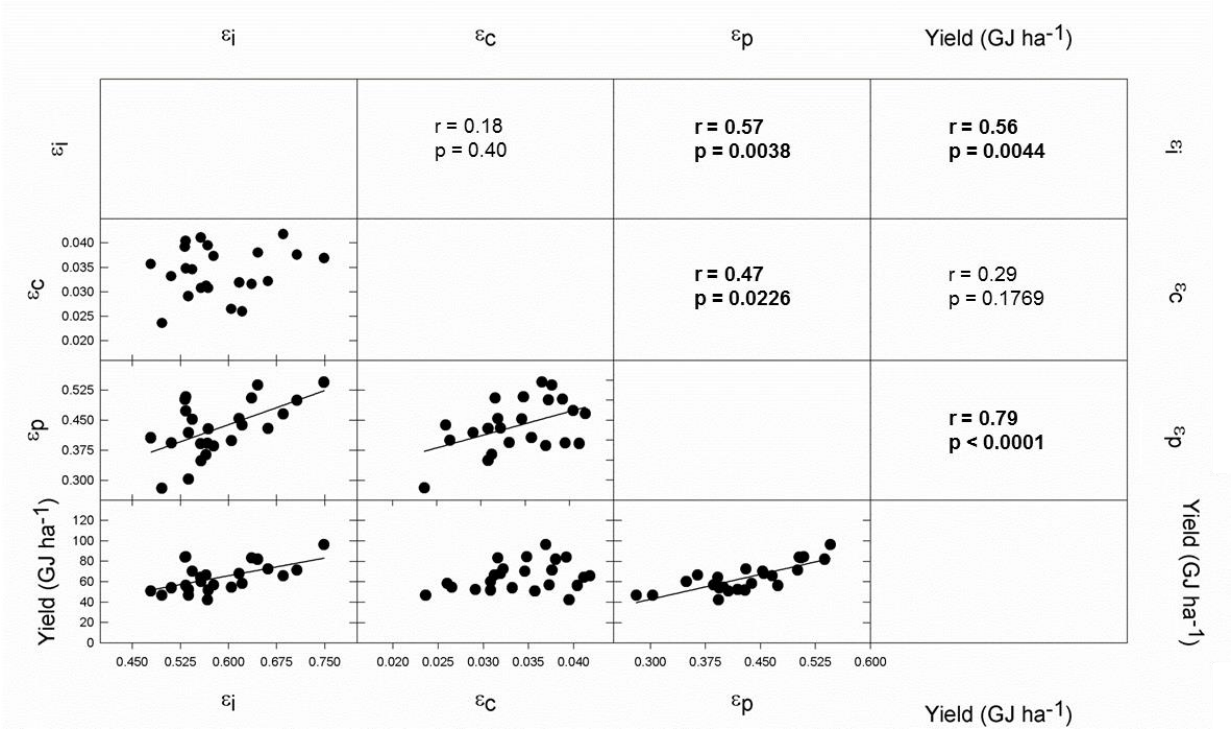

Figure 2.11. 2013 correlation matrix of yield and the Monteith efficiencies. Scatterplots and correlation coefficients are plotted in a matrix where lines represent the significant least squares regression. The Pearson correlation coefficients (r) and p-values (p) are reported and shown in bold when significant.

CHAPTER III: GREATER PHOTOSYNTHESIS OF MODERN SOYBEAN CULTIVARS DEPENDS ON GREATER STOMATAL CONDUCTANCE, NOT INCREASED CARBOXYLATION OR ELECTRON TRANSPORT CAPACITY.

Introduction

Soybean is an important source of protein for food and feed throughout the world and is second only to maize in planted area in the United States (FAOSTAT). Soybean seed yields do not appear to be stagnating in most regions (Ray *et al.*, 2012; Rowntree *et al.*, 2013), and in fact, the annual rate of both genetic and on-farm soybean yield improvement is now faster than it was 40 years ago (Rincker *et al.*, 2014; Specht *et al.*, 2014). However, the current rate of yield gain is still insufficient to meet growing demand and the United Nation's 2050 target of doubling crop yields (Ray *et al.*, 2013). Climate change further challenges yield progress as environmental stressors such as increased heat and drought negatively impact crop production (Lobell *et al.*, 2014; Ort and Long, 2014). Analysis of historical soybean germplasm has documented that breeders have increased yields by increasing plant harvest index (HI; i.e., the seed fraction of all aboveground biomass in terms of weight or energy), and to a lesser extent, seasonal canopy light interception and seasonal conversion efficiency (*CE*; i.e., the effective utilization of solar energy to produce plant biomass) (Koester *et al.*, 2014). The historical improvements in HI and canopy light interception efficiency have put these two physiological yield components close to their proposed limits (Hay, 1995; Zhu *et al.*, 2010). Soybean *CE*, however, is still below the estimated theoretical limit for oxygenic photosynthesis, and thus *CE* may become the next-generation target for sustaining future improvement in soybean productivity (Zhu *et al.*, 2010).

Recent reviews (Zhu *et al.*, 2010; Parry *et al.*, 2011; Raines, 2011) have proposed that improving *CE* through greater photosynthetic carbon fixation (*A*) could provide the bump in crop performance needed to sustain and improve yields for a growing population. While our previous work with historical soybean germplasm showed that long-term selection for ever-greater yield has improved *CE* (Koester *et al.*, 2014), the mechanistic basis for the enhancement in *CE* is not known. In Canadian and Chinese soybean germplasm, there is evidence that increased *A* has accompanied genetic yield improvement (Morrison *et al.*, 1999; Jin *et al.*, 2010), but a comprehensive study on how *A* has been affected by breeding in the US soybean germplasm is absent. In replicated field trials conducted over three successive years (2011-2013), we took over 4,000 measurements of leaf gas-exchange in 24 soybean cultivars of Midwestern US

adaptation that had year of release dates (YOR) spanning the period 1923 to 2007 (Table 3.1). We sought to understand the mechanisms driving the improvements in *CE* by investigating how photosynthetic and respiratory capacities, i.e., the investment in Rubisco protein activity, electron transport capacity, and mitochondrial electron transport, had been altered by 84 years of soybean genetic yield improvement.

Materials and Methods

Details of the experimental design for this study have been reported in the previous chapter with exception that photosynthesis data collected in 2011 was included in this analysis. In 2011, the experimental plots were not thinned to a uniform density, and so data from 2011 was not used to calculate light interception efficiency, conversion efficiency, partitioning efficiency or seed yield. Plots were thinned to a uniform density in 2012 and 2013. Daily meteorological data including S_t (Fig. 3.1A-C), temperature (Fig. 3.1D-F), and precipitation (Fig. 3.1G-I) were collected ~1.5 km from the field site by the Illinois Climate Network monitoring station (Angel *et al.,* 2009). Plots were irrigated using drip-line tubing four times during the 2012 season to relieve water stress (Fig. 3.1H). In 2013, a very wet spring prevented the deployment of the drip-line irrigation system.

Diurnal measurements of gas exchange were conducted on 14 days across the 2011-2013 growing seasons. For each diurnal, leaf CO_2 and water vapor exchange was measured approximately every 2-3 hours during the daytime. Two to three sunlit, fully expanded leaves from different plants in each plot were measured at each time-point. Each time-point was completed in 45 to 60 minutes. Leaf CO_2 and water vapor exchange were measured using infrared gas analyzers (LI-6400, LI-COR, Lincoln, NE), which were able to control temperature, the photosynthetic photon flux density (PPFD), CO_2, and relative humidity in the sample cuvette. Temperature and PPFD were held at ambient conditions measured immediately before each time-point and were kept constant throughout each time-point. Ambient PPFD was measured with a quantum line sensor (LP-80, LI-COR) and temperature was measured using the thermocouple within the cuvette of the infrared gas analyzer. The concentration of CO_2 was set to 400 ppm and relative humidity was adjusted to 60-65%. Leaf photosynthesis (A), stomatal conductance (g_s), and intercellular concentration of CO_2 (c_i) were calculated using the equations of von Caemmerer and Farquhar (1981). The total daily CO_2 uptake (A') was calculated from

the instantaneous *A* measurements during the diurnal samplings by summing the trapezoidal area under the curve. When dew was present on the leaves during the morning time-points, the leaves were dried before gas exchange was measured. During the time-points where dew was present, g_s was not calculated to avoid over estimates arising from residual moisture on the leaf surface.

In 2012 and 2013, the maximum rates of Rubisco carboxylation ($V_{c,max}$) and electron transport (J_{max}) were estimated at two stages of reproductive growth (flowering (R2) and pod fill (R5)) by measuring *A* at 12 CO_2 concentrations. Trifoliates were excised at the petiole from three plants per plot pre-dawn and were cut again under water and transported to the laboratory. The trifoliates remained in the dark until ~20 min before measurement when the leaves were illuminated to adapt to light conditions. Reference CO_2 concentration was initially set at 400 ppm and was reduced stepwise to 50 ppm. Thereafter the reference CO_2 concentration was restored to 400 ppm and increased stepwise to 2000 ppm as described in Ainsworth *et al.* (2002). The measurements were made at PPFD of 1750 μmol m^{-2} s^{-1} and at 25°C. Photosynthetic parameters were calculated by fitting the equations of Farquhar *et al.* (1980) and by creating an A/C_i response plot as described in Long and Bernacchi (2003). $V_{c,max}$ was fit using the points that fell below the inflection point of the A/c_i plot and J_{max} determined from the points above.

Nighttime rates of leaf respiratory CO_2 efflux were measured on two dates (flowering and pod fill) in the 2012 and 2013 growing seasons. Measurements were made on attached, mature leaves at the top of the canopy from 22:00-1:00 using a custom built chamber designed for the LI-6400 gas exchange system (LI-COR) as described in Gillespie *et al.* (2012). Full trifoliates were sealed in the chamber and were allowed to equilibrate for approximately 5-7 minutes until the relative humidity within the chamber was between 60-65% and steady rates of CO_2 efflux were obtained.

Soil volumetric water content was modeled using the Soil Temperature and Moisture Model (STM^2; USDA-ARS; Spokas and Forcella, 2009). The model predicts soil temperature and moisture conditions based on soil type, incident solar radiation, maximum and minimum temperatures, and precipitation. Weather parameters used in the model were collected from the Illinois Climate Network monitoring station, and soil moisture data were averaged for the top 30 cm of the soil profile.

Correlations between variables and cultivar YOR were tested for significance using least square regression (PROC MIXED, SAS version 9.2, SAS Institute Inc., Cary, NC, USA) or first

order linear regression (SigmaPlot, Systat Software, Inc, Richmond, CA, USA). A student's t-test was used to determine if average *A* and g_s values were significantly different. To select the model of best fit of the relationship between *A* and g_s during midday time-points (11:00-14:00 hrs), the Akaike information criterion (AIC) was calculated for a linear, quadratic, logarithmic, piecewise, and exponential rise to maximum model using PROC NLMIXED. The exponential rise to maximum model best fit the relationship by having the lowest AIC value. Correlation matrices were constructed using R (R Foundation for Statistical Computing, Vienna, Austria).

Results and Discussion

In this three year field study, integrated total daily carbon uptake (*A'*), increased linearly with cultivar YOR on approximately half of the soybean growth stage-specific measurement dates (Fig. 3.2). Recently released cultivars averaged 12% greater carbon uptake on a leaf area basis than older cultivars on all but one of those eight days. The exception was an even greater advantage on the last measurement date in 2013 (Fig. 3.2N), arising from much lower *A'* values in older cultivars because they had started to senesce by that sampling date. Modern cultivars had greater rates of *A* primarily in the early afternoon when *A* peaked (Figs. 3.3F-J, 3.4F-J, 3.5E-H). During these periods, modern cultivars had up to 23% greater rates of *A* than older cultivars. In Canadian and Chinese germplasm, slightly higher gains in leaf-level *A* with cultivar YOR were detected, although the absolute values of photosynthesis and seed yield were much lower (Morrison *et al.*, 1999; Jin *et al.*, 2010).

In C_3 plants under light-saturating conditions, photosynthetic rates can be limited by the kinetics of carbon fixation and substrate regeneration. Ribulose-1,5-bisphosphate (RuBP) carboxylase/oxygenase (Rubisco) is the primary carboxylating enzyme in soybean and other C_3 plants, and the rate at which Rubisco attaches CO_2 to its substrate, RuBP, can limit carbon gain (Farquhar *et al.*, 1980; Farquhar and Sharkey, 1982). When the maximum rate of carboxylation by Rubisco ($V_{c,max}$) is no longer limiting, the rate at which RuBP is regenerated (J_{max}) can become the limiting factor (Long *et al.*, 2006). Therefore improving $V_{c,max}$ and J_{max} have been identified as strategies to increase photosynthesis, biomass, and seed yield in crop species (Parry *et al.*, 2011; Raines, 2011; Long *et al.*, 2006). In the current study, neither $V_{c,max}$ nor J_{max} increased with cultivar YOR, were not correlated with either *CE* or seed yield (Figs. 3.6, 3.7), indicating that these two photosynthetic parameters were not changed during 84 years of

selection for greater yield. There was also no correlation with cultivar YOR and dark respiration rate (Figs. 3.6, 3.7). Similar results were reported for wheat for which neither $V_{c,max}$ nor J_{max} was correlated with biomass or yield (Driever *et al.*, 2014), and there was no evidence for increased dark respiration in historical wheat germplasm (Sadras and Lawson, 2011). The fact that gains in A' in modern soybean and wheat germplasm have not been achieved by improved maximum photosynthetic capacity begs the question as to why $V_{c,max}$ or J_{max} improvement did not co-occur with long term soybean yield improvement that is now averaging 29 kg ha^{-1} y^{-1} (Rincker *et al.*, 2014). Answering this question is the key to determining if photosynthetic capacity is truly an untapped target for increasing soybean productivity in the future.

Since A' was not increased by gains in maximum photosynthetic capacity ($V_{c,max}$ or J_{max}), then greater carbon fixation had to arise from enhanced stomatal conductance (g_s) for H_2O and CO_2. During diurnal time-points where significant increases in *A* were observed, proportional increases in g_s were detected (Fig. 3.3, 3.4, 3.5) suggesting that improvements in *A* with cultivar YOR were attained by the capacity to sustain greater CO_2 uptake and transpirational water use. Plotting and fitting midday *A* versus g_s for the 5 oldest cultivars and the 5 most modern cultivars using an exponential rise to a maximum model ($y = a(1 - e^{-bx})$) supports this interpretation (Fig. 3.8). When g_s was low (i.e., <0.4 mol m^{-2} s^{-1}), there was no difference in *A* between the older and more recently released cultivars (Fig. 3.8). However, when g_s was > 0.4 mol m^{-2} s^{-1}, there was a significantly greater *A* (~8%) in modern cultivars and there was also a significant positive trend in A' vs. YOR (Fig. 3.8). Similar results of increased g_s with years of breeding and seed yield have been found within the historical soybean (Morrison *et al.*, 1999) and wheat (Fischer *et al.*, 1998; Sadras and Lawson, 2011; Zheng *et al.*, 2011) germplasm, but the mechanisms behind greater g_s in modern cultivars remain unknown. It could be that modern soybean cultivars have greater stomatal sensitivity to environmental conditions or that they have greater maximum stomatal pore size or density. However, given that atmospheric CO_2 concentration increased from ~300 ppm in 1920 (Etheridge *et al.*, 1996) to 384 ppm in 2007 (NOAA website), and stomatal density has been shown to decrease with increasing atmospheric CO_2 concentration in many tree species (Woodward, 1987), we hypothesize the greater maximum conductance in modern soybean lines is caused by greater pore size.

The environmental conditions that the soybean canopy experienced largely determined the average g_s. During the three years of the study, the experimental plots experienced a wide

variety of environmental conditions including periods of drought and above average temperatures (Fig. 3.1). Modeled soil volumetric water content (SWC) estimated from the Soil Temperature and Moisture Model (STM^2; USDA-ARS; Spokas and Forcella, 2009), varied from 0.26 cm^3 H_2O cm^{-3}soil to 0.32 cm^3 H_2O cm^{-3} soil, and there was a quadratic relationship between average g_s across all cultivars and SWC (Fig. 3.9), as has been previously reported for soybean (Gilbert *et al.*, 2011a). STM^2 predicts SWC based on a general soil characterization and daily temperature and precipitation data. Therefore, it cannot distinguish between different rates of potential water use in the historical soybean cultivars studied. We hypothesize that more modern lines are able to extract more soil moisture, and thus would more rapidly deplete soil moisture reserves, although this hypothesis remains to be tested.

Conclusion

Recent research has shown that modern soybean cultivars can take advantage of high-yielding environments more than older cultivars, so historical rates of genetic gain are greater in better environments (Rincker *et al.*, 2014), and these data provide a physiological explanation for that. While modern cultivars have greater yields as a result of improved harvest index and greater light interception under both favorable and less favorable environments (Koester *et al.*, 2014; Rincker *et al.*, 2014), only under adequate soil moisture conditions do modern cultivars also have higher average g_s and *A*. Thus, in selecting for maximum seed yield, soybeans have been bred to take advantage of replete soil moisture conditions. If soil moisture content is reduced and drought becomes more prevalent with global climate change, then different strategies for improving *CE* in soybean will be required in the future.

Tables and Figures

Table 3.1. List of maturity group III soybean cultivars grown with respective year of release (YOR) dates and plant introduction (PI) number.

Cultivar	YOR	PI No.
Dunfield	1923	PI548318
Illini	1927	PI548348
AK (Harrow)	1928	PI548298
Mandell	1934	PI548381
Lincoln	1943	PI548362
Adams	1948	PI548502
Ford	1958	PI548562
Shelby	1958	PI548574
Ross	1960	PI548612
Adelphia	1964	PI548503
Wayne	1964	PI548628
Calland	1968	PI548527
Williams	1971	PI548631
Woodworth	1974	PI548632
Zane	1984	PI548634
Private 3- 2	1986	n/a[†]
Resnik	1987	PI534645
Private 3- 9	1989	n/a
Private 3-19	1994	n/a
Private 3-11	1996	n/a
IA 3010	1998	n/a
IA 3023	2003	n/a
Private 3-13	2004	n/a
Private 3-14	2007	n/a

[†]not available

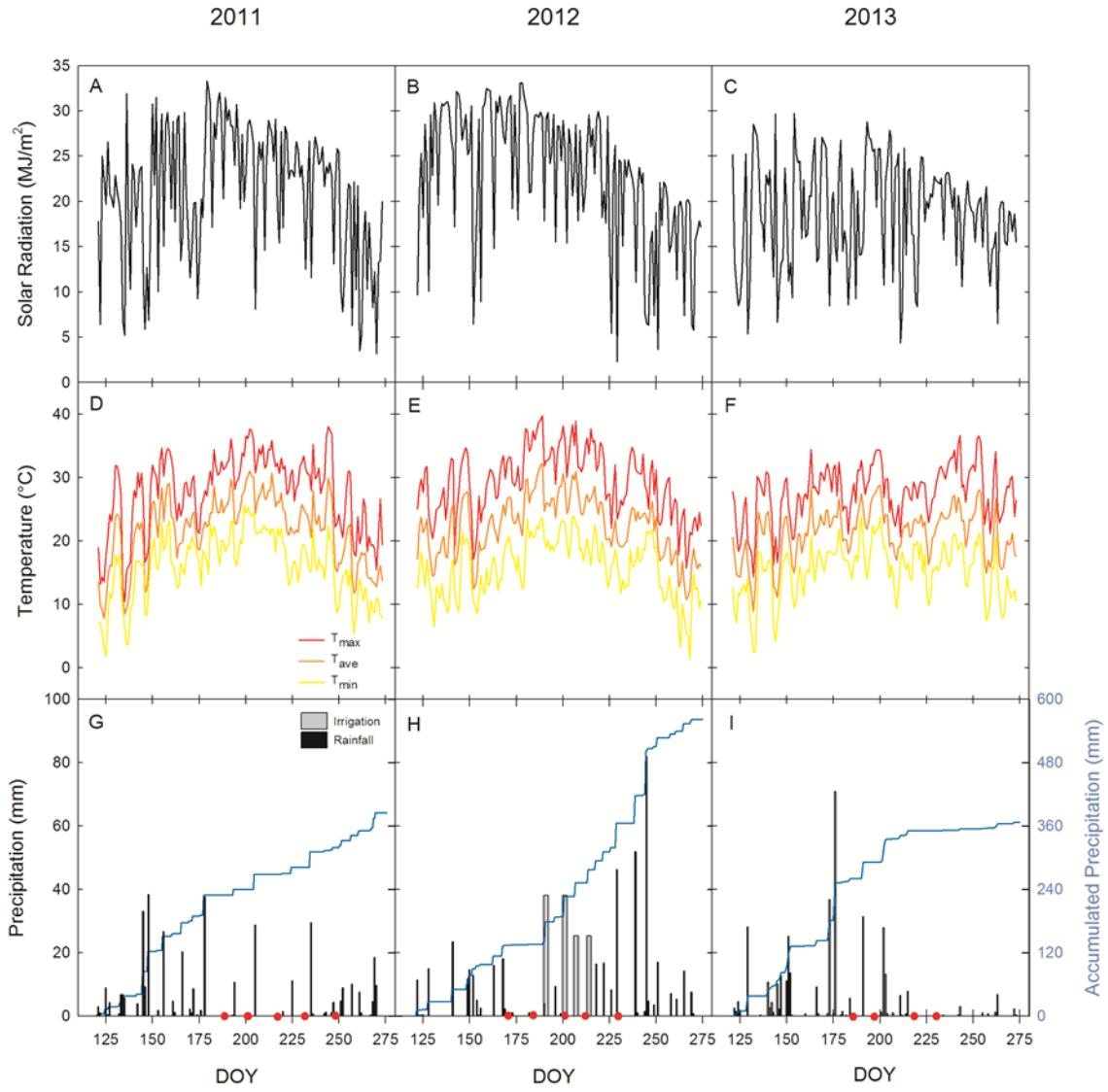

Figure 3.1. Meteorological data for the experimental growing seasons.
Meteorological conditions collected for the 2011-2013 growing seasons. Daily total solar radiation (A-C), daily maximum (red), average (orange), and minimum (yellow) temperatures (D-F), daily rainfall (black bars) and irrigation events (grey bars) and accumulated precipitation across the growing season (blue line) (G-I). Sampling dates for photosynthetic measurements are also shown with red circles (G-I).

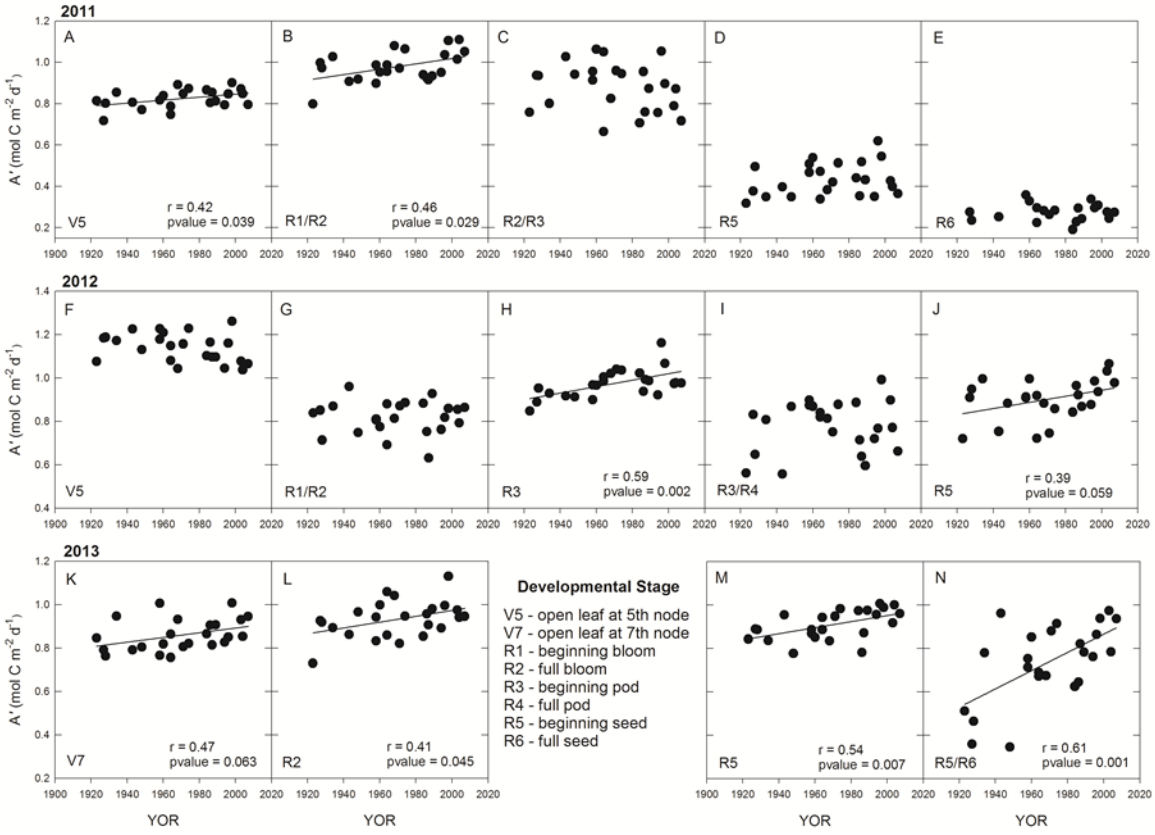

Figure 3.2. Daily carbon uptake versus cultivar year of release (YOR). (**A-E**) Diurnal air temperature (open triangles) and photosynthetic photon flux density (PPFD) (closed triangles). (**F-J**) Average assimilation rate (*A*) of the five oldest cultivars (closed circles) and five most recently released cultivars (open circles). (**K-O**) Average stomatal conductance (g_s) of the five oldest cultivars (closed circles) and five most recently released cultivars (open circles). The developmental stage of each sampling date is shown at the top of the plot (see legend in Fig. 1). Significant differences between old and new cultivars at $p < 0.05$ and $p < 0.01$ are denoted by * and **. † denotes when g_s data was not reported because leaves were damp.

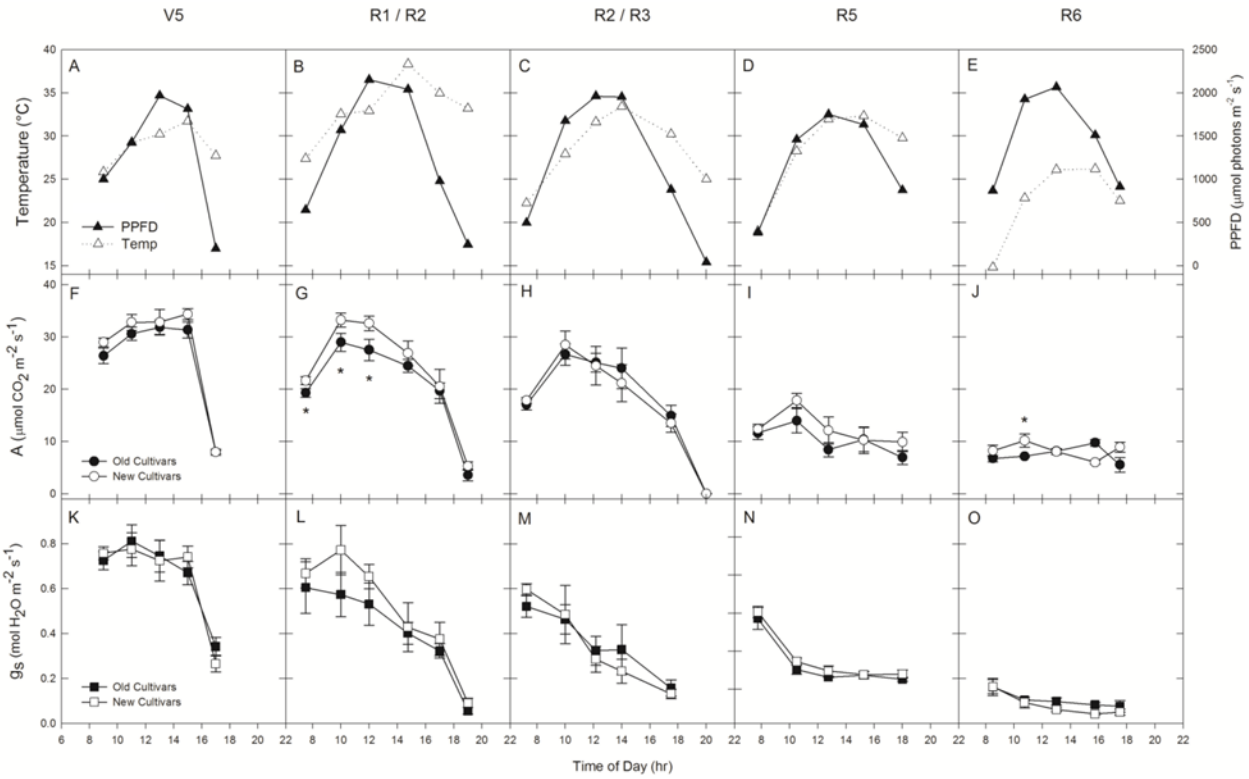

Figure 3.3. Diurnal photosynthetic rates and stomatal conductance during the 2011 growing season. Diurnal air temperature (open triangles) and photosynthetic photon flux density (PPFD) (closed triangles). (**F-J**) Average assimilation rate (*A*) of the five oldest cultivars (closed circles) and five most recently released cultivars (open circles). (**K-O**) Average stomatal conductance (g_s) of the five oldest cultivars (closed circles) and five most recently released cultivars (open circles). The developmental stage of each sampling date is shown at the top of the plot (see legend in Fig. 1). * $p < 0.05$

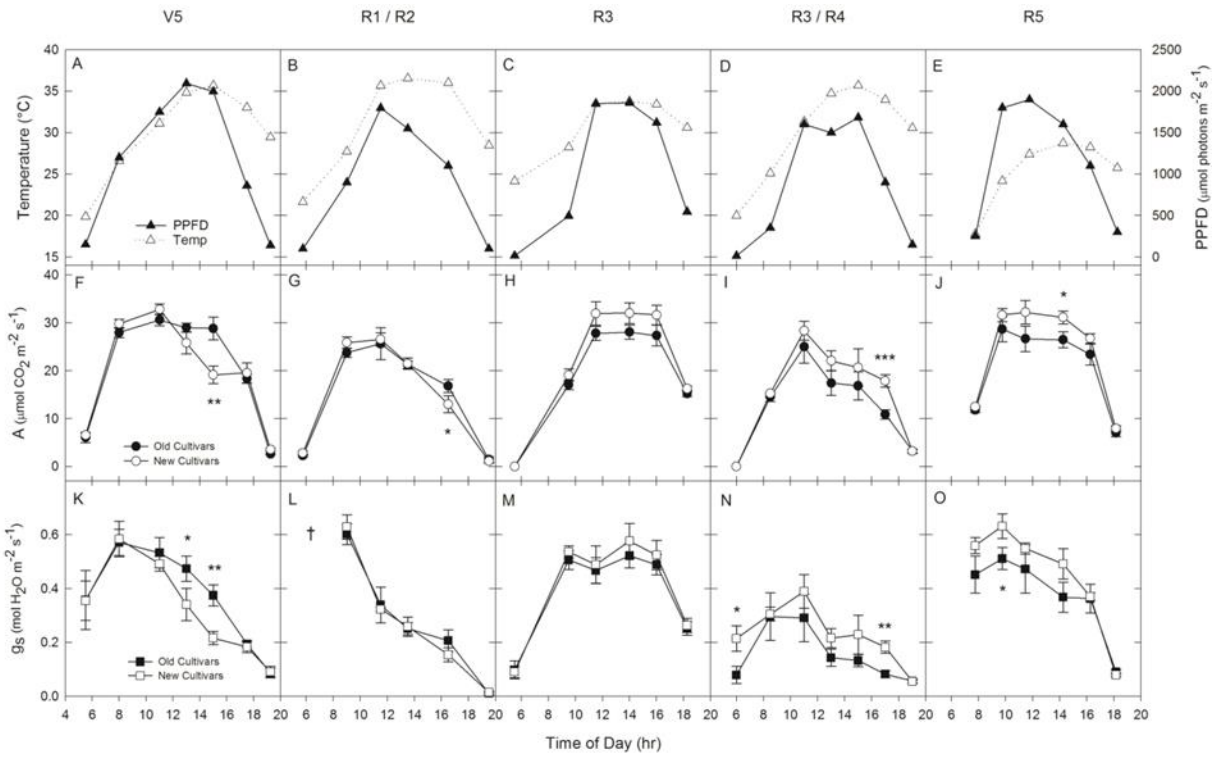

Figure 3.4. Diurnal photosynthetic rates and stomatal conductance during the 2012 growing season. (**A-E**) Diurnal air temperature (open triangles) and photosynthetic photon flux density (PPFD) (closed triangles). (**F-J**) Average assimilation rate (*A*) of the five oldest cultivars (closed circles) and five most recently released cultivars (open circles). (**K-O**) Average stomatal conductance (g_s) of the five oldest cultivars (closed circles) and five most recently released cultivars (open circles). The developmental stage of each sampling date is shown at the top of the plot (see legend in Fig. 1). Significant differences between old and new cultivars at $p < 0.05$ and $p < 0.01$ are denoted by * and **. † denotes when g_s data was not reported because leaves were damp.

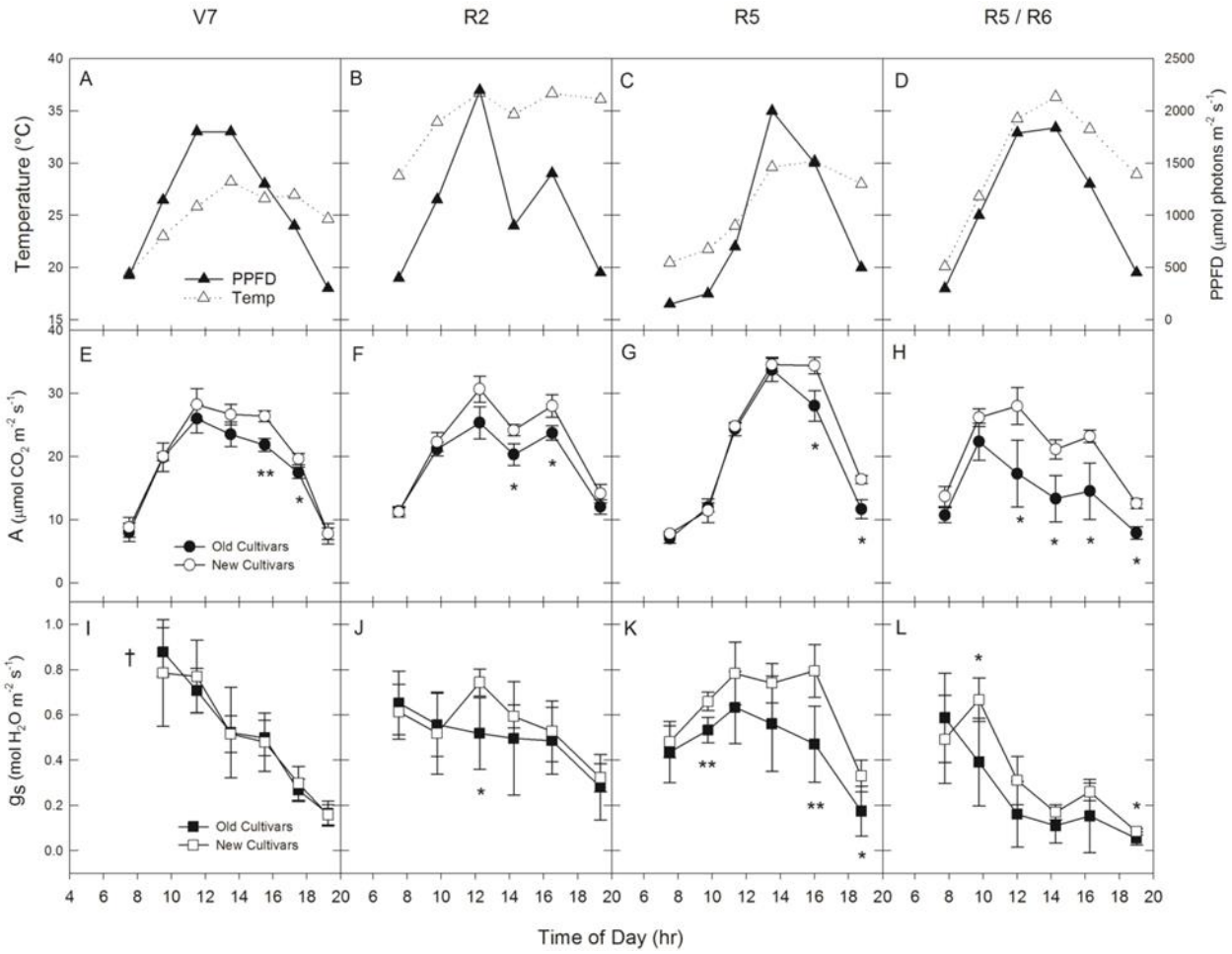

Figure 3.5. Diurnal photosynthetic rates and stomatal conductance during the 2013 growing season. Diurnal air temperature (open triangles) and photosynthetic photon flux density (PPFD) (closed triangles). (**F-J**) Average assimilation rate (*A*) of the five oldest cultivars (closed circles) and five most recently released cultivars (open circles). (**K-O**) Average stomatal conductance (g_s) of the five oldest cultivars (closed circles) and five most recently released cultivars (open circles). The developmental stage of each sampling date is shown at the top of the plot (see legend in Fig. 1). * $p < 0.05$, ** $p < 0.01$

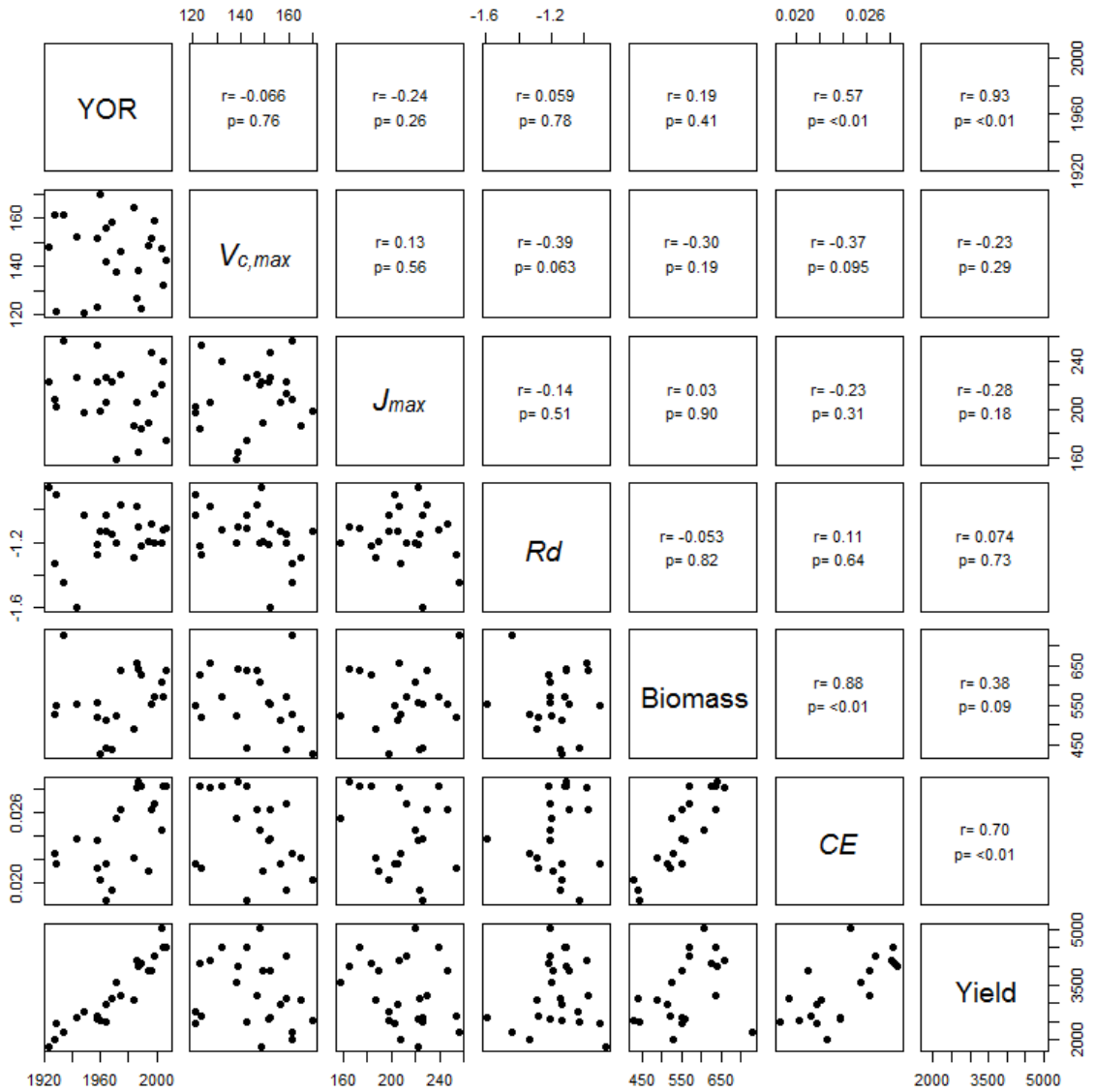

Figure 3.6. Correlation matrix of cultivar YOR, photosynthetic traits, respiration, and yield in 2012. The correlation matrix of cultivar YOR, photosynthetic traits, respiration, and yield is shown for the 2012 growing season. The Pearson coefficient (r) and the p-value (p) are reported.

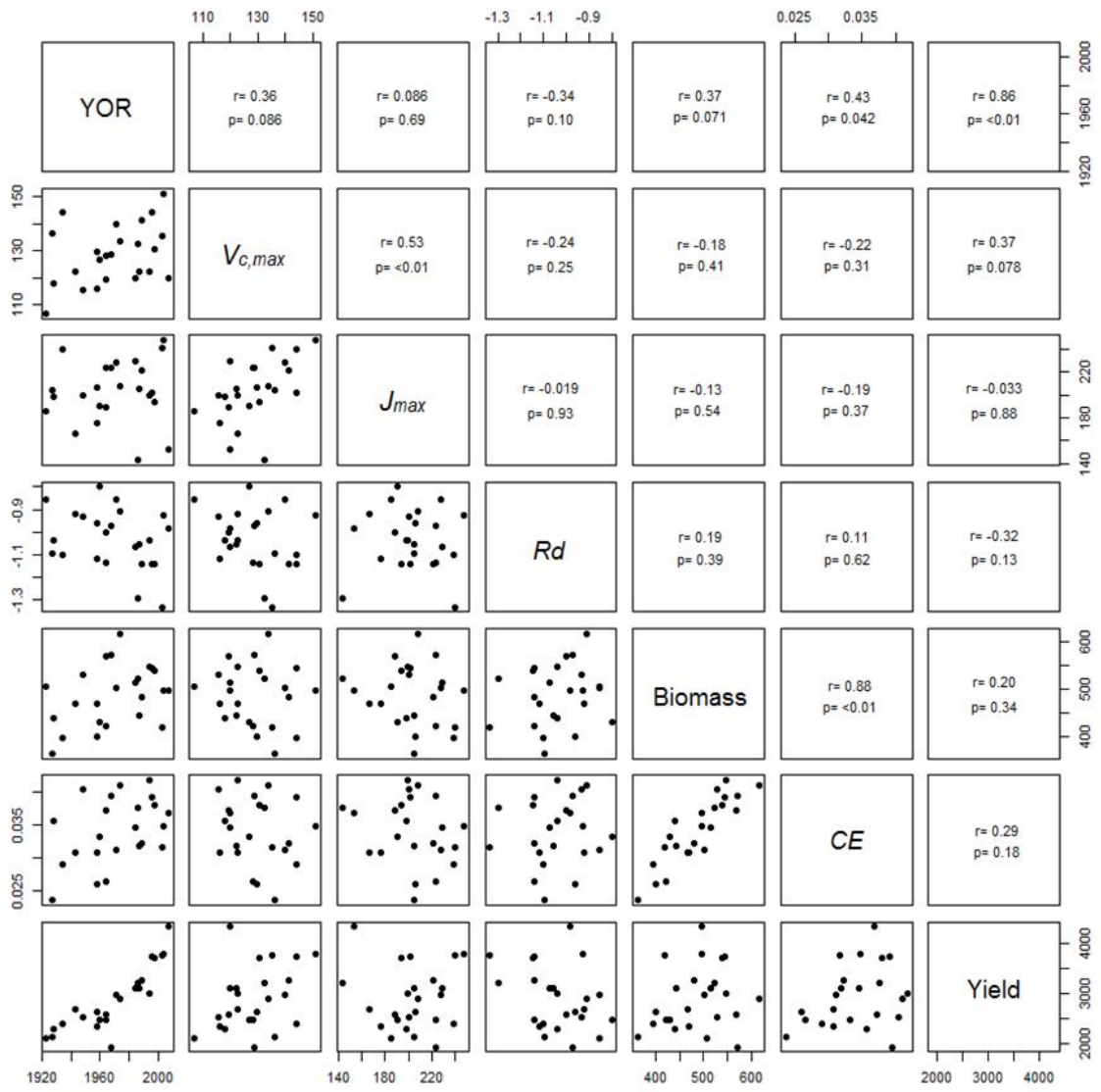

Figure 3.7. Correlation matrix of cultivar YOR, photosynthetic traits, respiration, and yield in 2013. The correlation matrix of cultivar YOR, photosynthetic traits, respiration, and yield is shown for the 2013 growing season. The Pearson coefficient (r) and the p-value (p) are reported.

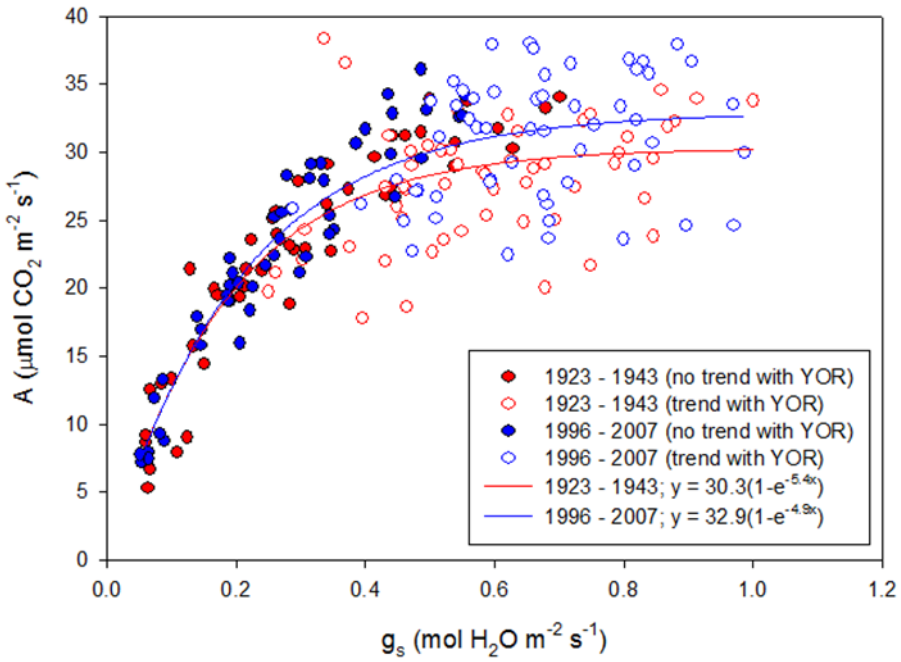

Figure 3.8. Relationship between photosynthetic rate and stomatal conductance. The relationship between photosynthetic rate (A) and the stomatal conductance (g_s) in the five oldest (red) and newest cultivars (blue) across all measurement dates is shown. Closed circles are timepoints were no significant trend between A' and YOR was found and open circles represent timepoints were there was a linear relationship between A' and YOR.

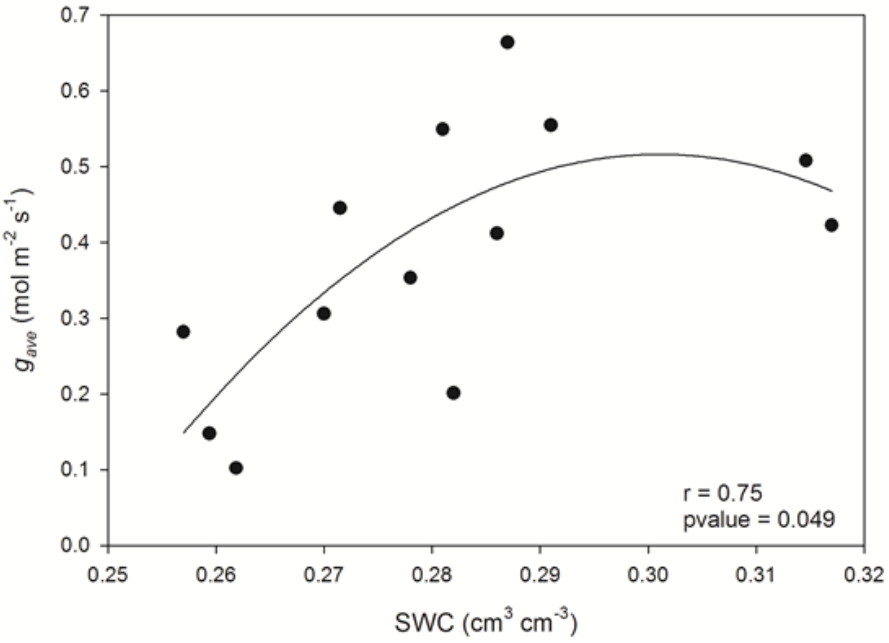

Figure 3.9. Relationship between stomatal conductance and soil moisture content. The quadratic function between the average stomatal conductance across all soybean lines (g_{ave}) and the soil volumetric water content (SWC) on each of the sampling dates. The Pearson correlation coefficient (r) and p-value are reported.

CHAPTER IV: ALTERED TRANSCRIPT ABUNDANCE OF RUBISCO ACTIVASE AND VEGETATIVE STORAGE PROTEINS CORRELATE WITH YIELD INCREASES IN HISTORIC SOYBEAN GERMPLASM

Introduction

Soybean (*Glycine max*) is the most widely grown leguminous crop (FAOSTAT, 2013) and is a major source of protein and oil for food, feed, and fuel. Yields have steadily increased over the past century due to advancements made by breeding and management strategies (Specht *et al.*, 1999; Rowntree *et al.*, 2013). However, with world population expected to reach 9 billion by the end of the century, the current yield trends may not be sufficient to supply the demands of a burgeoning population (Tilman *et al.*, 2011; Ray *et al*., 2013). Past gains in yield have been driven by greater allocation of carbon to reproductive structures, enhanced light interception by the canopy, and improved energy conversion efficiency (ε_c; Koester *et al.*, 2014), but little is known about specific changes in gene expression underpinning these improvements.

One potential avenue of improving yield gains is to transform so-called "yield enhancement genes" (YEG) into crop species. YEG are single genes, that when altered, are hypothesized to increase plant growth and yield (Van Camp, 2005; Gonzalez *et al.*, 2009). Of course, yield is a complex, quantitative trait, and analyses have demonstrated that yield is controlled by many genes across numerous developmental and physiological processes (Orf *et al.*, 1999; Guzman *et al.*, 2007; Palomeque *et al.*, 2009). Nevertheless, there are multiple proposed strategies of increasing crop production through genetic transformation and a large majority aim to enhance processes of photosynthesis, protein modification and synthesis, growth and development, or source-sink relations (Van Camp, 2005; Busov *et al.*, 2008; Gonzalez *et al.*, 2009; Ainsworth *et al.*, 2012). Support for these strategies for yield improvement comes from studies showing that overexpression of single genes can increase biomass and/or yield, albeit in model species or under artificial growth conditions (Choe *et al.*, 2001; Deprost *et al.*, 2007; Gonzalez *et al.*, 2009; Preuss *et al.*, 2012; de Bossoreille de Ribou *et al.*, 2013). Additionally, compelling correlations between gene expression and yield have been reported (Yin *et al.*, 2010; Preuss *et al.*, 2012), indicating that single genes can have an impact on yield. While some of the target genes have begun to be tested in crop species in the field (Yin *et al.*, 2010, Preuss *et al.*, 2012), most still require multi-location field trials across multiple years in important agricultural crops to verify whether these YEG manipulations translate to greater agronomic production.

Although there has been a considerable amount of research on historical yield improvement, there is little known about the genetic control underlying past yield determination (Ross-Ibarra *et al.*, 2007). By examining potential genetic drivers of past yield advancement, the genetic mechanisms underlying yield production may be better understood and could point to promising targets for further improvements. Therefore, the objective of this study is to examine how the expression of 15 putative YEG correlate with yield within the context of the historical soybean germplasm. Soybean cultivars from the maturity group III with year of release dates (YOR) spanning 1923-2007 (Table 1) were used to test in a common-garden experiment if the expression of YEG has changed with traditional breeding. Target genes were chosen based on published relationships between expression and biomass and/or yield in either *Arabidopsis thaliana* or soybean (Table 2).

Methods

Experimental design and tissue sampling

Soybean cultivars were grown in the field as previously described by Koester *et al.* (2014). Leaf tissue for quantitative PCR was harvested during two stages of reproductive development (full flowering, R2 and beginning seed, R5) in 2012 and 2013. Five leaflets from five different plants per cultivar were sampled between 12:00 and 13:00, placed in a foil packet and immediately frozen in liquid N. Tissue was stored at -80 °C.

RNA extraction and preparation

All 5 leaflets from individual cultivars were pooled and ground to a fine powder in liquid N using a mortar and pestle. Total RNA was extracted from the ground leaf tissue using the PureLink Plant RNA Reagent (Ambion, Life Technologies Corporation, Grand Island, NY, USA) according to the manufacturer's protocol. RNA was quantified with a spectrophotometer (NanoDrop 1000, Thermo Fischer Scientific, Waltham, MA, USA), and RNA quality was assessed with the Agilent 2100 Bioanalyzer (Agilent Technologies, Santa Clara, CA, USA). Genomic DNA contamination was removed from RNA samples using the DNA-free DNase Treatment and Removal kit (Applied Biosystems/Ambion, Austin, TX, USA) according to the manufacturer's protocol. First-strand complementary DNA (cDNA) was created using Superscipt II RNase H- Reverse Transcriptase (Life Technologies, Grand Island, NY, USA) in a

20 µL reaction volume using 1 µg of the DNase treated RNA and oligo(dT) primers according to the manufacturer's instructions.

Quantitative PCR

qPCR was performed on an Applied Biosystems 7900HT Fast Real-Time PCR system (Life Technologies) using the Power SYBR Green PCR master mix (Life Technologies) and 400 nM of each primer in a 10 µL total reaction volume as described in Yendrek *et al.* (2012). Primer sequences for each target gene and reference gene are described in Table 4.3. Target gene expression was normalized to cons8 (Libault *et al.*, 2008) which is a Peptidase S10, serine carboxypeptidase (Fig. 4.1). PCR amplification curves were constructed and analyzed using the LinRegPCR software (Ruijter *et al.*, 2009) which determined the PCR efficiency and the critical threshold value from the baseline-corrected delta-Rn values in the log-linear phase. These values were then used to calculate the relative expression of each gene as described in Gillespie *et al.* (2008) and were then normalized to the average expression for that gene across all of the cultivars examined.

Yield, conversion efficiency, and protein yield

The total seed yield and ε_c values used in the correlation matrix were taken from Koester *et al.*, (2014). Seed samples were measured for protein content using a Perten DA 7200 Feed Analyzer (Perten Instruments, Stockholm, Sweden). To determine protein yield, the percent protein per seed was multiplied by total seed yield per plot.

Statistical analysis

A significant correlation between gene expression and cultivar YOR was tested using least squares regressions (PROC REG procedure, SAS version 9.2, SAS Institute Inc., Cary, NC, USA) or first order linear regression (SigmaPlot, Systat Software, Inc, Richmond, CA, USA). Outliers were detected using PROC UNIVARIATE (SAS version 9.2, SAS Institute Inc., Cary, NC, USA) by testing if the residuals determined from linear regression fell outside the 95% confidence interval. The Cook' D test in PROC REG was used to identify points that were overly influential in determining the regression coefficients. Outliers and points that were overly influential were removed and data were re-normalized to the average expression of the remaining

data points. Correlation matrices were constructed using R (R Foundation for Statistical Computing, Vienna, Austria).

Results and Discussion

In field trials of U.S. soybean cultivars with YOR dates from 1923 to 2007, the expression of some YEG were shown to increase with cultivar YOR and to correlate with ε_c and yield. Specifically, three genes encoding Rubisco activase showed increased abundance in more recently released cultivars and had significant associations with ε_c and yield. Further, two genes encoding vegetative storage proteins had increased expression with cultivar YOR and strong correlations to yield gains. In the following sections, we describe each of the tested YEG and their correlation with historical yield gains in soybean.

Photosynthesis

Photosynthesis determines the supply of assimilate that is available for growth and biomass accumulation and therefore plays a key role in determining crop productivity (Gifford *et al.*, 1984, Parry *et al.*, 2011). In C_3 plants such as soybean, photosynthesis is often limited by the initial step in carbon fixation where ribulose-1,5-bisphophate (RuBP) carboxylase/oxygenase (Rubisco) fixes CO_2 to RuBP (Farquhar *et al.*, 1980). Owing to Rubisco's intimate role in photosynthetic carbon gain, increasing the amount of Rubisco or its activation state would theoretically increase yields (Parry *et al.*, 2011; Raines, 2011; Parry *et al.*, 2013). Rubisco activase is an AAA+ protein that uses energy from ATP to remodel Rubisco complexes and release its inhibitors, thereby maintaining Rubisco's catalytic function (Portis, 2003; Mueller-Cajar *et al.*, 2014). There is evidence that Rubisco activase transcript abundance correlates with soybean photosynthetic capacity and seed yield (Yin *et al.*, 2009). There are five genes in soybean that encode Rubisco activase (Yin *et al.*, 2009), but two of the genes' expression are not abundant in leaf tissue (http://soybase.org/soyseq/; Severin *et al.*, 2010). Therefore, we focused our analysis on the three other genes whose transcripts are more abundant in soybean leaf tissue (Table 2). The expression of three genes encoding Rubisco activase (*RCAα*, *RCAβ*, and *RCA-14*) was positively correlated with cultivar year of release (Fig 4.2). In 2012, both the expression of *RCAβ* and *RCA-14* significantly increased with cultivar year of release (Fig 4.2C, E), although *RCAα* expression did not (Fig. 4.2A). In 2013, the expression of *RCAα* and *RCAβ* was

significantly correlated with cultivar YOR, and *RCA-14* was weakly correlated with YOR (Fig. 4.2). In 2012, the increases in gene expression *RCAβ* and *RCA-14* were also significantly correlated with both the ε_c and final yield (Fig. 4.3), while in 2013, expression of all three activase genes was positively correlated with yield (Fig 4.4). These results support the previous research linking the expression of Rubisco activase to soybean yield (Yin *et al*., 2009).

Growth and development

Overexpression of the TARGET OF RAPAMYCIN (TOR) kinase increases *Arabidopsis* growth and seed production, while the inhibition of TOR kinase results in arrested plant growth (Deprost *et al.*, 2007). TOR kinase is a serine/threonine kinase that regulates transcription and translation of ribosomal components, which are important for cell growth (Wullschleger *et al.*, 2006). Within our study, two genes that encode the equivalent TOR kinase in soybean (*TOR-1* and *TOR-2*) were identified. *TOR-1* transcript abundance was too low to quantify with qPCR, but *TOR-2* significantly increased with years of traditional breeding (Fig. 4.6) and was weakly positively correlated with yield (Fig. 4.8) in 2013, but not 2012 (Figs. 4,5, 4.7). In our analysis of *TOR-2* abundance in historical soybean genotypes, many of the genotypes were outliers and so could not be used in the analysis. Thus, the results warrant further investigation. Still, the mechanism for the increase in seed production in historical soybean lines and in transgenic Arabidopsis over-expressing TOR is similar. In *Arabidopsis*, TOR kinase influences seed production by increasing the number of total siliques without altering seed or silique size (Deprost *et al.*, 2007). Similarly, historical soybean breeding has decreased vegetative growth and increased the number of seeds per plant (Rowntree *et al.*, 2013; Koester *et al.*, 2014), with little impact on individual seed mass (Morrison *et al*., 2000; Rowntree *et al.*, 2013).

Another mechanism by which increased plant growth and seed production has been increased in *Arabidopsis* is through altered synthesis of the hormone brassinolide by overexpressing an intermediate enzyme, DWARF4, involved in its synthesis (Choe *et al.*, 2001). Brassinolide is a brassinosteroid that mediates integral processes involved in cell division and elongation (Clouse and Sasse, 1998). The soybean homologues for the DWARF4 gene (*DWF-1* and *DWF-2*) were too low in abundance to measure using qPCR. This result is in line with literature that found that RNA levels of these transcripts in *Arabidopsis* were extremely low (Choe *et al.*, 1998), even though analysis of the expression profiles of putative DWARF4 genes

in the RNA-Seq Atlas (http://soybase.org/soyseq/; Severin *et al.*, 2010) found moderate levels of transcript abundance.

In a large-scale screen of candidate genes for yield improvement, the *Arabidopsis* BBX32 gene was identified as increasing yields in soybean (Preuss *et al.*, 2012). BBX32 encodes a B-box family protein that is related to the regulation of light signal transduction in *Arabidopsis* (Holtan *et al.*, 2011), and was found to regulate circadian clock genes in soybean (Preuss *et al.*, 2012). Further, gene homologues in soybean, *BBX52* and *BBX53*, increased seed yield when overexpressed in transgenic plants and tested in multi-locational field trials (Preuss *et al.*, 2012). Preuss *et al.* (2012) also found that the reproductive development period (R3-R7; Fehr *et al.*, 1971) was lengthened in the overexpression lines extending the period of pod and seed development and delaying senescence. In the current study, *BBX52* and *BBX53* did not correlate with cultivar YOR (Figs. 4.5, 4.6) or yield (Figs. 4.7, 4.8) in either year. The two genes, however, were consistently correlated with each other across years (Figs 4.7, 4.8). Within this set of historical lines, senescence (R7) among newer cultivars was delayed up to ~12 days (Koester *et al.*, 2014) which is much longer than the 3-4 day lengthening found in Preuss *et al.* (2012). Because there was no change in the expression of either *BBX52* or *BBX53*, it is likely that the later senescence found in these historical varieties is due to other mechanisms than through this protein.

Source-sink relations

Plants must regulate carbon metabolism to ensure there is adequate carbon assimilation and storage to supply current and future sinks, including the seed (Smith and Stitt, 2007). Therefore, altering carbon metabolism may improve source-sink relations leading to increased yields (Van Camp, 2005; Ainsworth and Bush, 2011; Ainsworth *et al.*, 2012). In elevated CO_2 studies of soybean, the chloroplast membrane triose phosphate translocator exhibited 33% greater expression at elevated CO_2 accompanying the stimulation in carbon gain and respiration (Leakey *et al.*, 2009). The triose phosphate translocator transports triose phosphates out of the chloroplast in exchange for free phosphate and is key in the regulation of many carbon metabolism networks (Stitt *et al.*, 2010). Transcript abundance of three genes encoding the triose phosphate translocator (*TPT-1*, *TPT-2*, *TPT-3*) was measured in the historical soybean lines, and there was no alteration of gene expression with cultivar YOR (Figs. 4.5, 4.6). *TPT-2*

was not abundant enough to detect transcript levels by qPCR. Further, only *TPT-1* correlated with yield in 2013, and was never associated with improvements in ε_c (Figs. 4.7, 4.8), suggesting that TPTs were not affected by historical soybean breeding.

Another candidate YEG, identified from global transcript profiling of 21 Arabidopsis accessions, is a Kelch repeat F-box protein (At1g23390), whose expression was positively correlated with plant fresh weight (Sulpice *et al.*, 2009). The soybean homolog encoding the Kelch repeat F-box protein (*FBX*) did not change with cultivar YOR across both years (Figs. 4.5, 4.6) and did not correlate with yield in either year (Figs. 4.7, 4.8). In addition to expression, allelic variation in *FBX* in Arabidopsis was also present and associated with biomass production (Sulpice *et al*., 2009); however, there does not appear to be any change in expression of FBX with year of release in soybean.

Finally, genes encoding two vegetative storage proteins (VSPα and VSPβ) were analyzed. Vegetative storage proteins (VSPs) are glycoproteins that accumulate in soybean paraveinal mesophyll and are hypothesized to facilitate the transfer and storage of carbon and nitrogen to leaf veins (Franceschi and Giaquinta, 1983). They are synthesized in sink organs, and then preferentially degraded as the organs transition to sinks (Staswick, 1989). In 2012, there was no change in the expression of the two genes (*VSPα* and *VSPβ*) encoding VSPs with cultivar YOR (Fig. 4.9A, C). However, in 2013, both *VSPα* and *VSPβ* showed increased expression with cultivar YOR and newer cultivars had approximately 2.5-fold greater expression than the average expression of the genes across all YOR (Fig. 4.9B, D). Both *VSPα* and *VSPβ* correlated with seed yield and seed protein yield in 2013 (Fig. 4.4) as would be expected if these proteins are playing a significant role at the onset of high sink demand. Contrary to the expected role of VSPs in nitrogen storage, down-regulation of *VSPα* and *VSPβ* in transgenic soybean had no effect on seed yield or seed protein content (Staswick *et al.*, 2001). However, the loss of VSPs in the transgenic plants was accompanied by an increase in other proteins, especially vegetative lipoxygenases (VLXs). VLXs also accumulate in the vacuoles of paraveinal mesophyll, which provide both storage and proteolytic function consistent with the proposed role in assimilate storage and mobilization (Murphy *et al*., 2005).

Conclusion

In conclusion, single genes hypothesized to be important targets for yield improvement were

tested in historical soybean cultivars. A positive correlation among some of the target genes and seed yield provides further support for the potential for over-expression of these YEG to improve soybean yields. Changes in YEG transcript abundance with cultivar YOR also provides insight into the genetics underlying past yield improvement. However, more work is needed to further understand the mechanisms of altered YEG transcript abundance and the functional implications for altered transcript abundance.

Tables and Figures

Table 4.1. List of maturity group III soybean cultivars grown with respective year of release (YOR) dates and plant introduction (PI) number. PI numbers from private lines were not available (n/a).

Cultivar	YOR	PI No.
Dunfield	1923	PI548318
Illini	1927	PI548348
AK (Harrow)	1928	PI548298
Mandell	1934	PI548381
Lincoln	1943	PI548362
Ford	1958	PI548562
Adelphia	1964	PI548503
Calland	1968	PI548527
Williams	1971	PI548631
Resnik	1987	PI534645
Private 3-11	1996	n/a
IA 3010	1998	n/a
Private 3-13	2004	n/a
Private 3-14	2007	n/a

Table 4.2. List of gene targets, associated gene function, and connection to increased biomass/yield determination.

Gene Name	**Function**	**Connection to Increased Biomass/Yield**
Kelch repeat F-box protein	unknown	positively correlated with biomass in *Arabidopsis* (Sulpice *et al.* 2009)
Triose phosphate: phosphate transporter	transporter of gylcerate-3-phosphate and orthophosphate	increased expression in elevated CO_2 (Leakey *et al.* 2009)
B-box domain gene	putative regulator of circadian clock genes	constitutive expression increased yield in soybean (Preuss *et al.* 2012)
TOR kinase	regulates transcription and translation of ribosomal components	overexpression increased growth and seed yield (Deprost *et al.*2007)
DWARF4	enzyme in brassinosteroid biosynthetic pathway	overexpression in Arabidopsis increased biomass and yield (Choe *et al.* 2001)
Rubisco activase	catalyzes the activation of Rubisco	eQTL analysis in soybean positively correlated gene expression and yield (Yin *et al.* 2010)
Vegetative storage protein	glycoprotein, storage of C and N	storage of C and N in soybean leaves (Franceschi and Giaquinta, 1983)

Table 4.3. List of genes, gene annotations, and the primers used for qPCR analysis.

Gene Name	Gene Abbreviation	Gene Annotation	Forward Primer	Reverse Primer
B-box domain gene	*BBX53*	Glyma07g02320	CCCTTCCTCTTCTACCTGCG	ATTTCTCAAACACCTCCGCC
B-box domain gene	*BBX52*	Glyma08g23700	TCCGCTGACTACTCCGATTC	TCAAACACCTCCTCCGATCC
DWARF4	*DWF4-1*	Glyma01g38180	TGGAGCAACACATAGCAAGG	ACCCGCATGTCTCTATGCAT
DWARF4	*DWF4-2*	Glyma11g07240	ATCCCAGAAGCATTGGTGGA	TGCTTCATCTTGGGCTGAGA
Kelch repeat F-box protein	*FBX*	Glyma01g03500	CAGTCGCCGGAGAAATCATG	GCCTTTCACGTTCCCAACTT
Rubisco activase	*RCAα*	Glyma02g41700	AGACGACCAGCAAGACATCA	TTGGGCAGGGTCATGAAGTT
Rubisco activase	*RCAβ*	Glyma11g34230	TCTCAGACGACCAACAGGAC	GGGCAGGGTCATGAAGTTCT
Rubisco activase	*RCA-14*	Glyma14g07270.1	CAAAGGTACCGTGAAGCTGC	TGTACATGCCAGGAAGCTGA
TOR kinase	*TOR-1*	Glyma01g45220	CCAACCATTCAAGGCCGATT	TCAAGTAGTTCGTGGCCCTT
TOR kinase	*TOR-2*	Glyma11g00480	TCGATGCTATCTGGGTTGCT	AACGGCAAGCAATGAACCAT
Triose phosphate transporter	*TPT-1*	Glyma04g35730	ACCCGGTAACTGTGACTGTT	AGCCACCTTTCCAAGACTCA
Triose phosphate transporter	*TPT-2*	Glyma06g19250	GGCCTCTTCCTTCTTCACCT	CAACGGCAAACTGAACAACA
Triose phosphate transporter	*TPT-3*	Glyma19g00270	ATCTCCATCCCAGACCAAGC	ACAAGGGAAGAAACCACCCA
Vegetative storage protein	*VSPα*	Glyma07g01730	GGAATACATCCATGGCGAAC	TTGCCCTTGTTAACCCATTC
Vegetative storage protein	*VSPβ*	Glyma08g21410	CGAAGTCCATCACAACGACA	TGAAGCCAAGAGACAACAGC

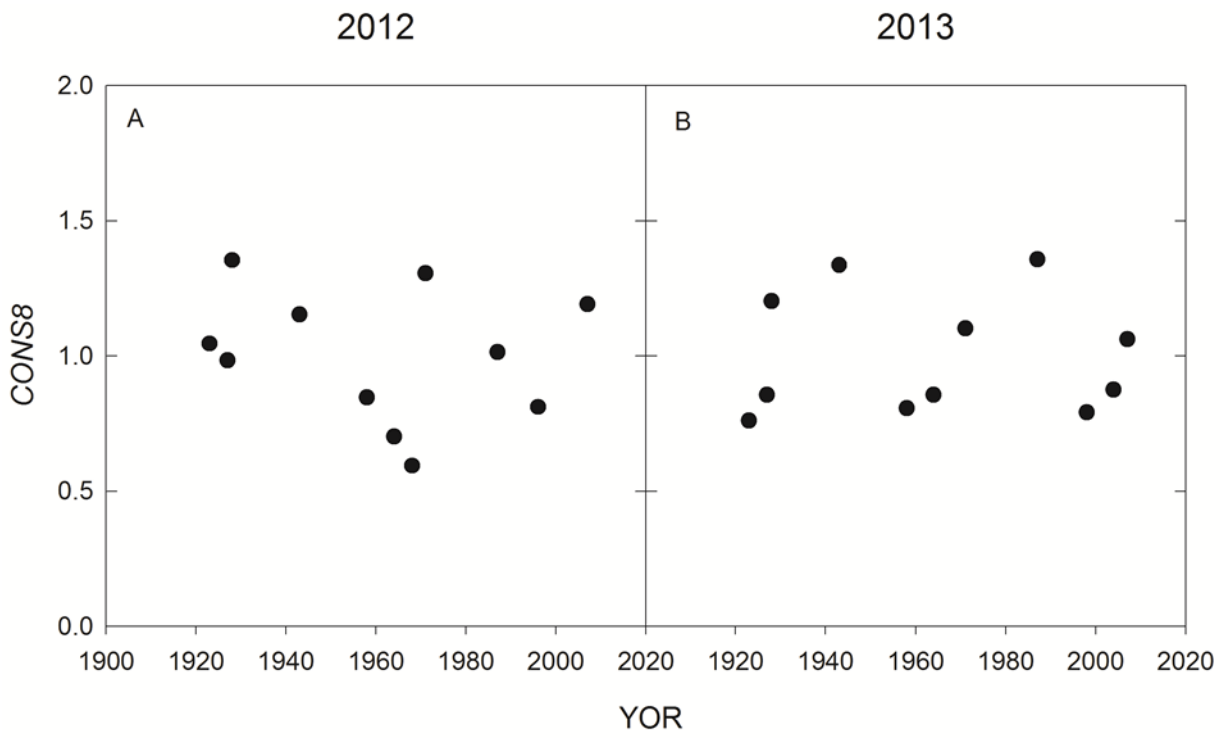

Figure 4.1. Correlation of reference gene expression with cultivar YOR in 2012 and 2013. Expression of the *CONS8* gene with cultivar YOR**.**

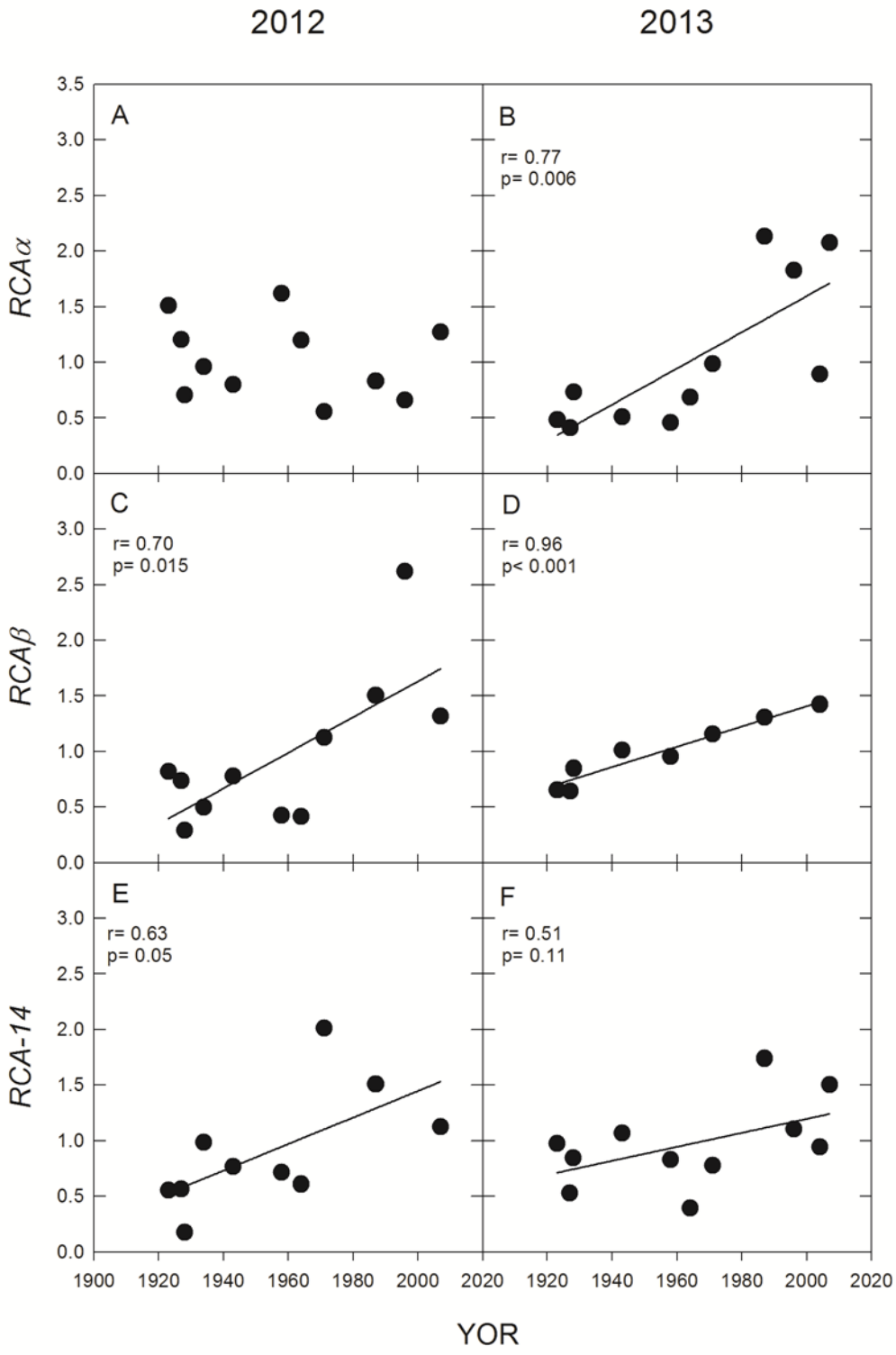

Figure 4.2. The correlation of genes encoding Rubisco activase with cultivar YOR during the 2012 and 2013 growing seasons. Three genes encoding Rubisco activase are plotted against cultivar YOR in the 2012 and 2013 growing seasons. The line is the linear least-squares regression between each variable. The Pearson's coefficient (r) and the p-value (p) is shown.

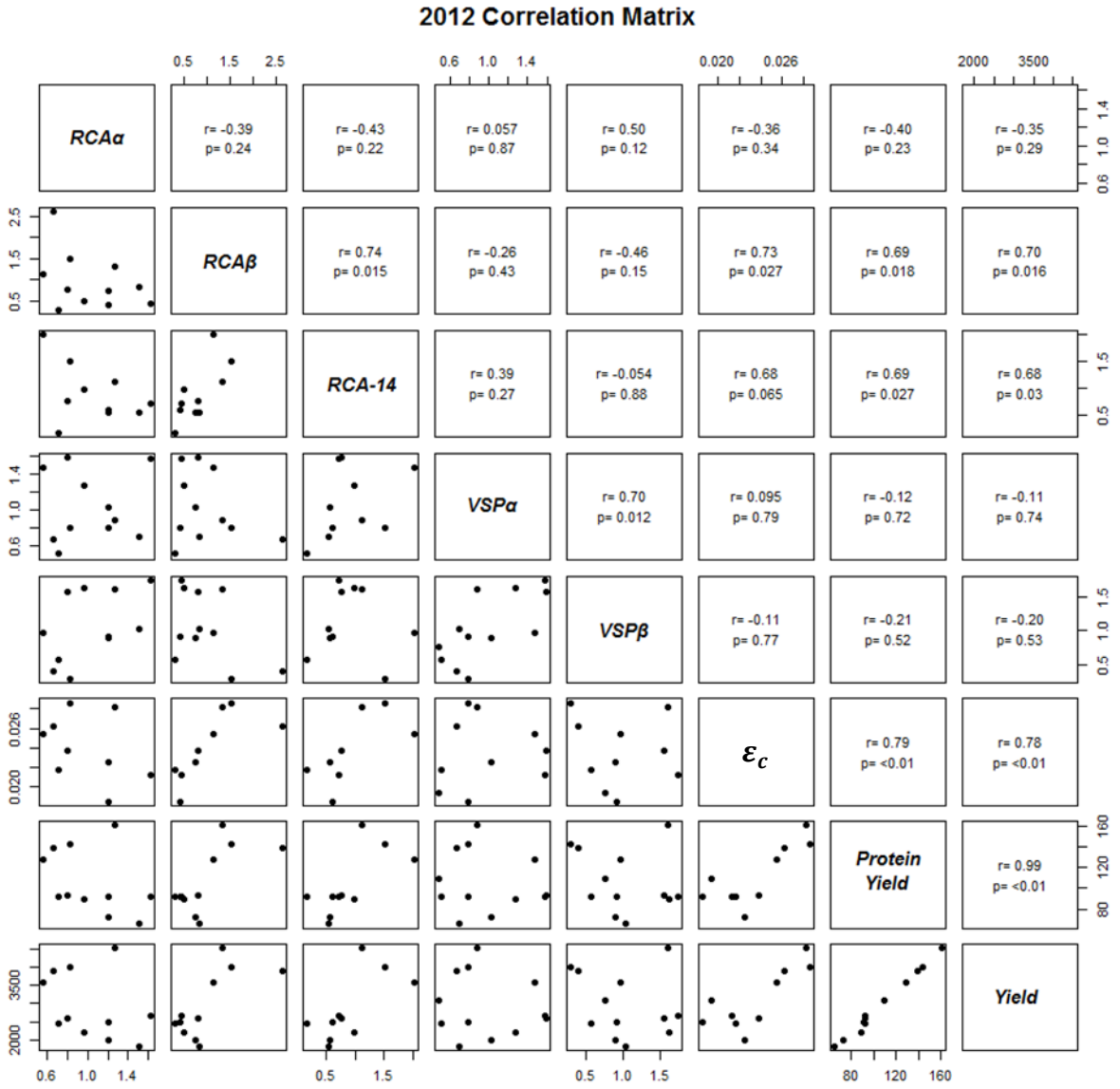

Figure 4.3. The correlation matrix of *RCAα, RCAβ, RCA-14, VSPα*, and *VSPβ* gene expression, ε_c, protein yield, and yield in 2012. Shown is the correlation matrix for the relative gene expression of *RCAα*, *RCAβ*, *RCA-14*, *VSPα*, and *VSPβ*, conversion efficiency (ε_c), protein yield, and yield in 2012. The top-right panels show the Pearson's coefficient (r) and the p-value (p) for each correlation. The bottom-left panels show the graphical correlation between each variable. ε_c and yield are taken from Koester *et al.*, 2014.

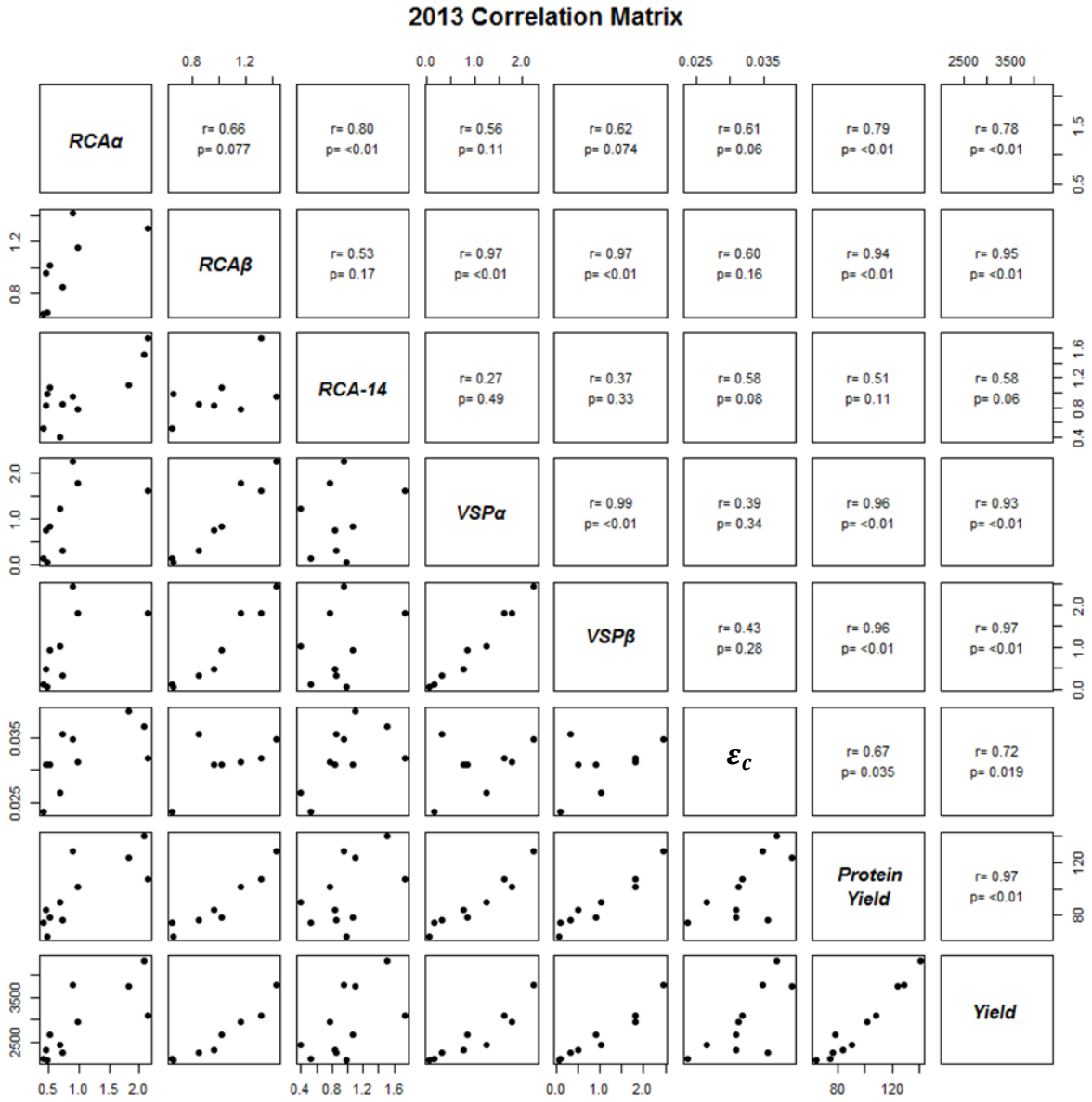

Figure 4.4. The correlation matrix of *RCAα, RCAβ, RCA-14, VSPα,* and *VSPβ* gene expression, ε_c, protein yield, and yield in 2013. Shown is the correlation matrix for the relative gene expression of *RCAα*, *RCAβ*, *RCA-14*, *VSPα*, and *VSPβ*, conversion efficiency (ε_c), protein yield, and yield in 2013. The top-right panels show the Pearson's coefficient (r) and the p-value (p) for each correlation. The bottom-left panels show the graphical correlation between each variable. ε_c and yield are taken from Koester *et al.*, 2014.

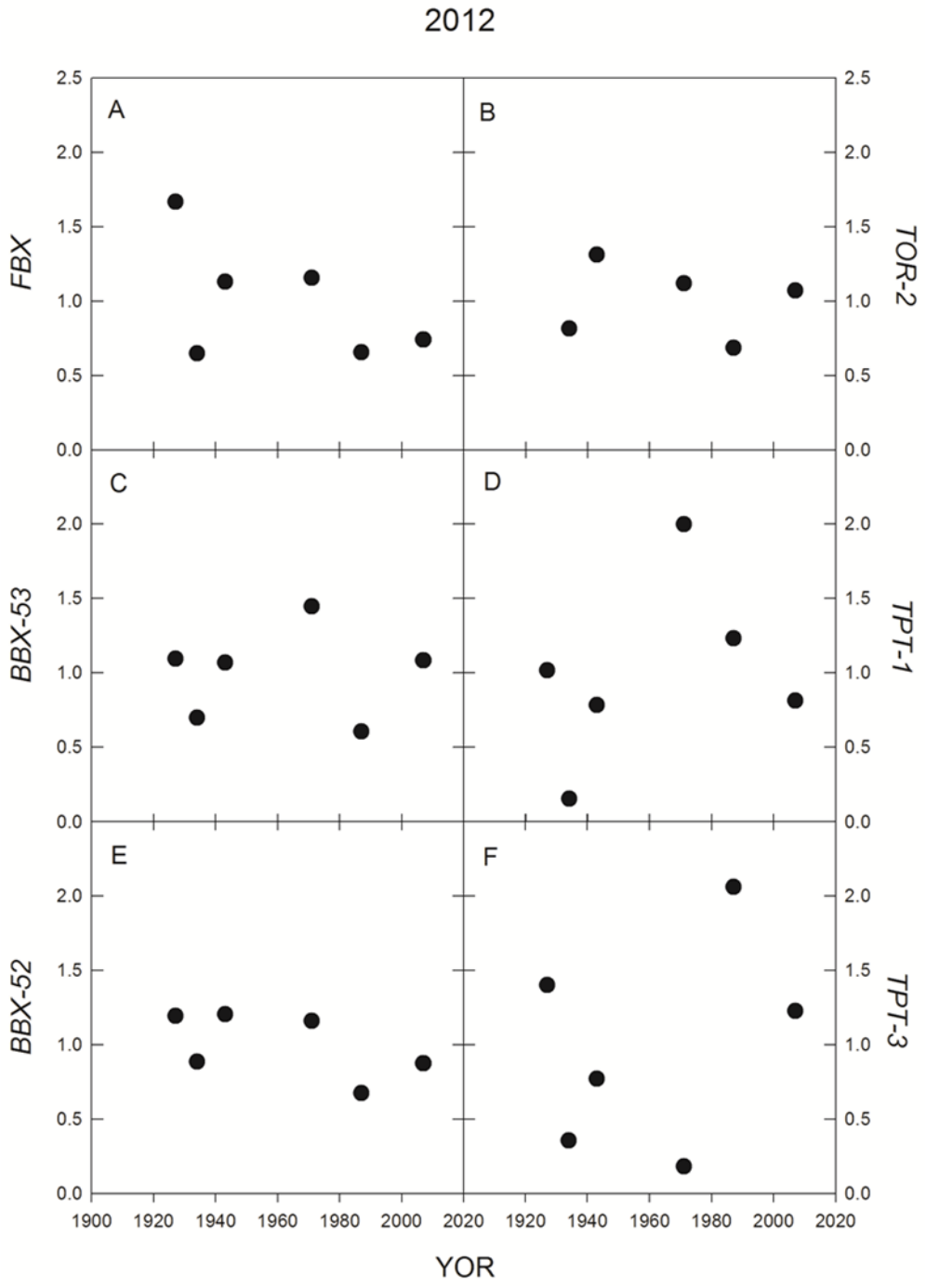

Figure 4.5. The *FBX*, *BBX52*, *BBX53*, *TOR-2*, *TPT-1*, and *TPT-3* gene expression do not change with cultivar YOR in 2012. The relative gene expression for the *FBX*, *BBX52*, *BBX53*, *TOR-2*, *TPT-1*, and *TPT-3* genes are shown plotted with cultivar YOR during the 2012 growing season.

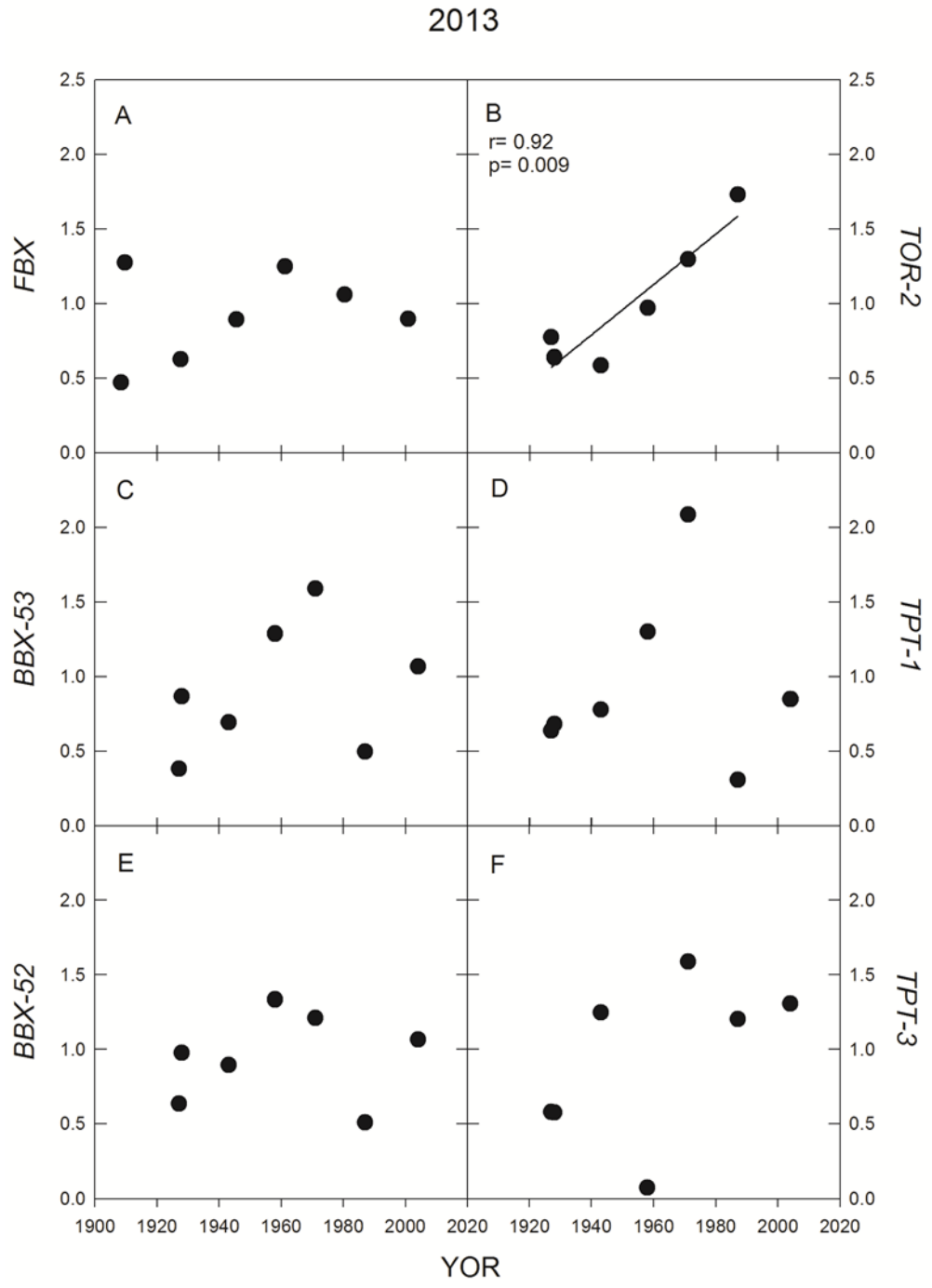

Figure 4.6. The *FBX*, *BBX52*, *BBX53*, *TOR-2*, *TPT-1*, and *TPT-3* gene expression do not change with cultivar YOR in 2013. The relative gene expression for the *FBX*, *BBX52*, *BBX53*, *TOR-2*, *TPT-1*, and *TPT-3* genes are shown plotted with cultivar YOR during the 2013 growing season.

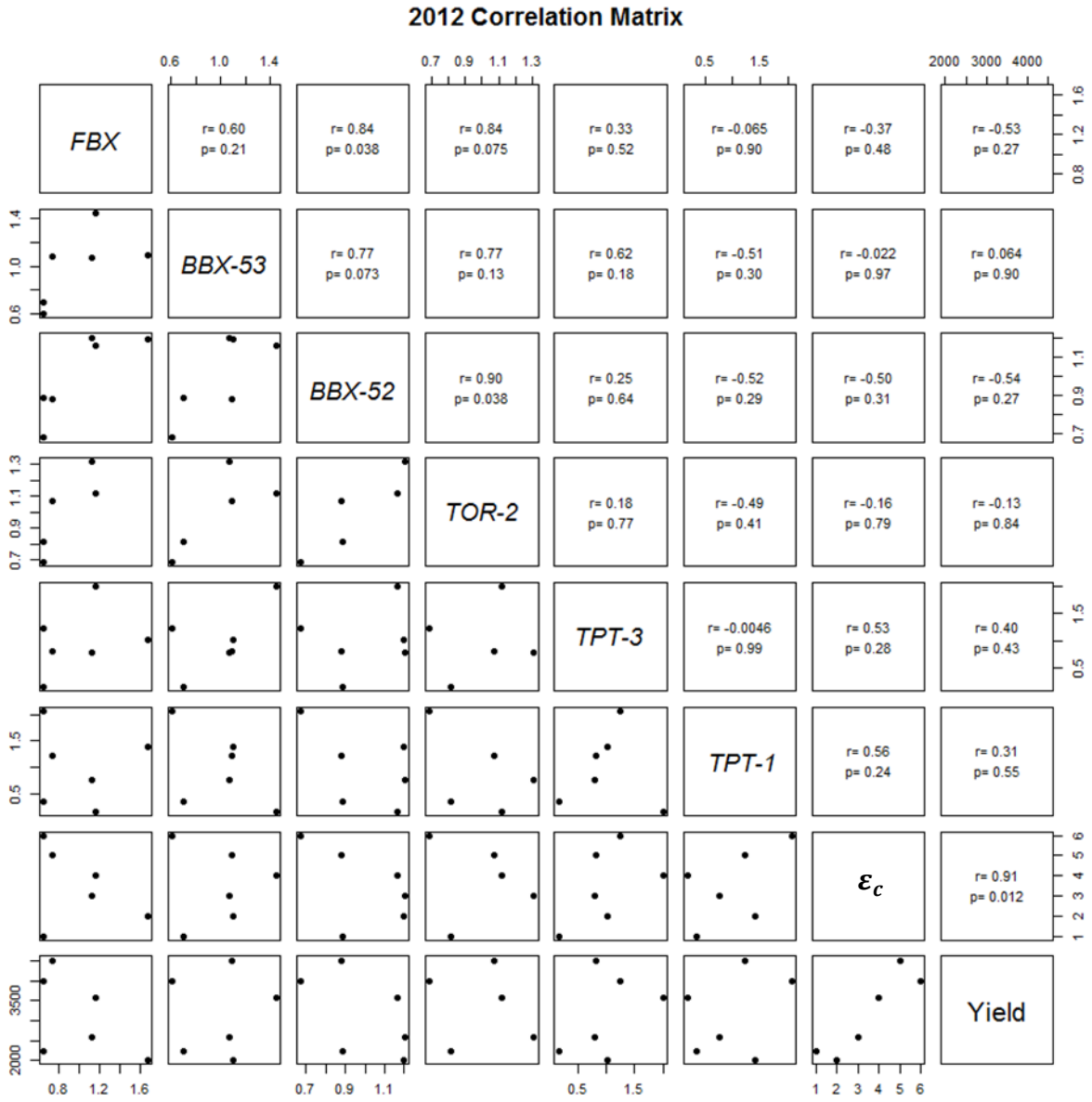

Figure 4.7. The correlation matrix of *FBX*, *BBX52*, *BBX53*, *TOR-2*, *TPT-1*, and *TPT-3* gene expression, ε_c and yield in 2012. Shown is the correlation matrix for the relative gene expression of *FBX*, *BBX52*, *BBX53*, *TOR-2*, *TPT-1*, and *TPT-3*, conversion efficiency (ε_c), and yield in 2012. The top-right panels show the Pearson's coefficient (r) and the p-value (p) for each correlation. The bottom-left panels show the graphical correlation between each variable. ε_c and yield are taken from Koester *et al.*, 2014.

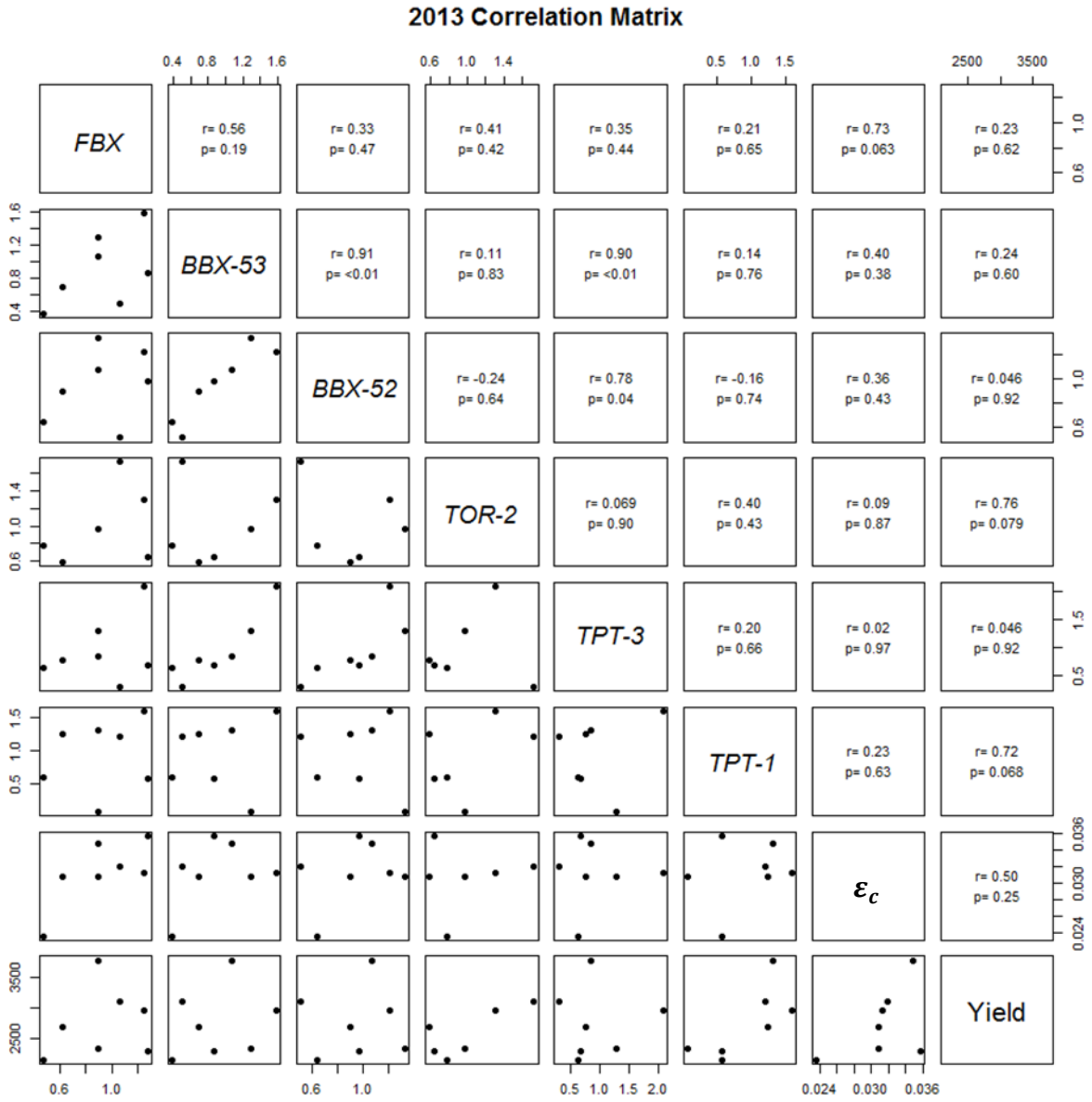

Figure 4.8. The correlation matrix of *FBX*, *BBX52*, *BBX53*, *TOR-2*, *TPT-1*, and *TPT-3* gene expression, ε_c and yield in 2013. Shown is the correlation matrix for the relative gene expression of *FBX*, *BBX52*, *BBX53*, *TOR-2*, *TPT-1*, and *TPT-3*, conversion efficiency (ε_c), and yield in 2013. The top-right panels show the Pearson's coefficient (r) and the p-value (p) for each correlation. The bottom-left panels show the graphical correlation between each variable. ε_c and yield are taken from Koester *et al.*, 2014.

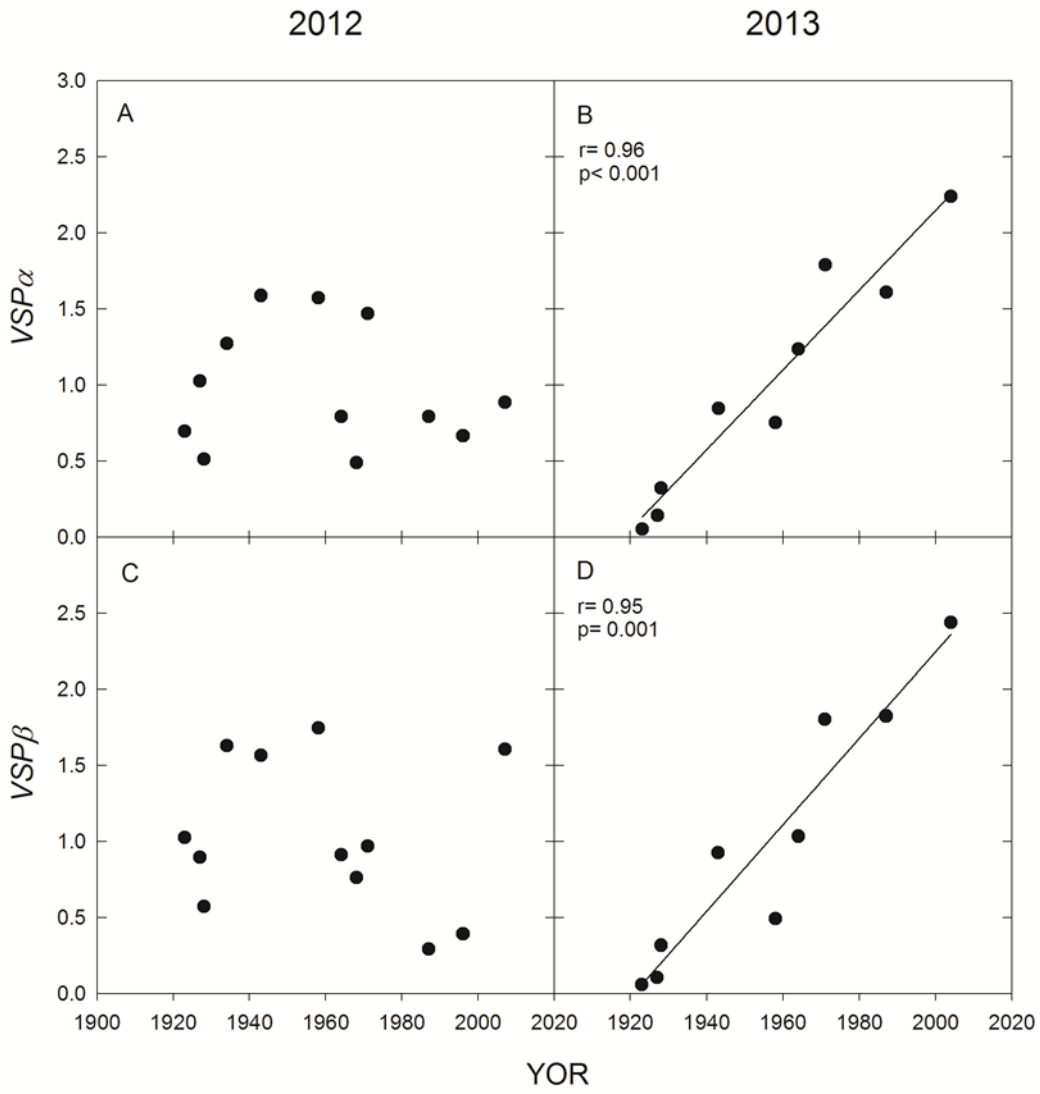

Figure 4.9. The correlation of genes encoding two vegetative storage proteins with cultivar YOR during the 2012 and 2013 growing seasons. Two genes encoding vegetative storage proteins are plotted against cultivar YOR in the 2012 and 2013 growing seasons. The line is the linear least-squares regression between each variable. The Pearson's coefficient (r) and the p-value (p) is shown.

CHAPTER V: CONCLUDING REMARKS

Despite the impressive gains made in soybean yields over the past century, the current pace of crop improvement is projected to fall short of the 2050 target of doubling crop yield (Ray *et al.*, 2013). The physiological basis for seed yield improvements in historical soybean cultivars was examined in this thesis. What has this retrospective analysis taught us about the prospects of enhancing future yield? Chapter II demonstrated that the efficiencies of light use (ε_i) and utilization (ε_c) along with harvest index (ε_p) were important in achieving the gains in yield made by traditional breeding. Although it appears that neither ε_i or ε_p are beginning to plateau, modern cultivars are nearing the maximum values that have been predicted for these parameters (Hay, 1995; Zhu *et al.* 2010), and therefore, there may be little room for further improvement of these efficiencies. However, the amount of solar energy intercepted by a soybean canopy could be increased with earlier planting date, more rapid canopy closure, and lengthening of the growing season (Zhu *et al.*, 2010; Parry *et al.*, 2011; Rowntree *et al.*, 2013, 2014). While ε_c has been increased through traditional breeding, it is still below the theoretical maximum suggesting it is an important target for future crop improvement (Zhu *et al*., 2010). However, greater absolute values of ε_c do not always result in greater seed yields in soybean, and do not consistently correlate with yield across different years of study (Chapter II). Thus, more research is needed to better understand how ε_c is influenced by the environment (Slattery *et al.*, 2014) and how changes in ε_c impact seed yield to guide future breeding strategies.

Chapter III provided evidence that the improvements in ε_c were driven by gains in photosynthetic daily carbon gain, and that the increases in photosynthesis were sustained through greater stomatal conductance. Because greater carbon acquisition came at the expense of increased water use, improvements in photosynthesis were only apparent during periods of ample water availability. This is important as a majority of soybeans are grown under water-limiting conditions (Boyer, 1982) and irrigation of this acreage is unsustainable. The data from Chapter III support the need to improve water-use efficiency (WUE) in soybean in order to maintain production in the future. Strategies for improving WUE in crops include breeding for high yielding cultivars that maintain a low transpiration rate (Condon *et al.*, 2004; Parry *et al.*, 2005) and identifying cultivars that achieve constant transpiration rates at high vapor pressure deficits (Gilbert *et al.*, 2011a). While some of these strategies have been successfully

incorporated into wheat (Rebetzke *et al.*, 2002), improvements in WUE do not always lead to greater yields and are often associated with decreases in photosynthesis (Condon *et al.*, 2004; Gilbert *et al.*, 2011b). Because of the intimate connection between photosynthesis and water use, it is imperative that strategies to improve ε_c are in the context of a warmer future.

In Chapter IV, I investigated the correlation between transcript abundance of putative yield enhancement genes (YEG) and yield within historical germplasm. While several of these YEG had correlations with yield, none of these proposed YEG fell in genic regions identified to be QTL for seed yield (http://soybase.org). This discrepancy perhaps illustrates that seed yield is a complex, multi-genic trait, and research is only beginning to understand the underlying genetic mechanisms governing yield formation. However, the continually growing genetic resources in soybean including the development of a Nested Association Mapping population in soybean should improve the ability to understand the genetic architecture of complex quantitative traits like seed yield (Stupar and Specht, 2013).

REFERENCES

Ainsworth EA, Rogers A, Nelson R, Long SP. 2004. Testing the 'source-sink' hypothesis of down-regulation of photosynthesis in elevated CO_2 in the field with single gene substitutions in *Glycine max*. *Agricultural and Forest Meteorology* **122,** 85-94.

Ainsworth EA, Bush DR. 2011. Carbohydrate export from the leaf: a highly regulated process and target to enhance photosynthesis and productivity. *Plant Physiology* **155,** 64-69.

Ainsworth EA, Davey PA, Hymus GJ, Drake BG, Long SP. 2002. Response of photosynthesis to elevated carbon dioxide in Florida scrub-oak ecosystem. *Ecological Applications* **12,** 1267-1275.

Ainsworth EA, Long SP. 2005. What have we learned from 15 years of free-air CO_2 enrichment (FACE)? A meta-analytic review of the responses of photosynthesis, canopy properties and plant production to rising CO_2. *New Phytologist* **165,** 351-372.

Ainsworth EA, Yendrek CR, Skoneczka JA, Long SP. 2012. Accelerating yield potential in soybean: Potential targets for biotechnological improvement. *Plant, Cell and Environment* **35**, 38-52.

Akita A. 1988. Physiological basis of heterosis in rice. *Hybrid Rice.* International Rice Research Insititute, Los Banos, Phillipines.

Amthor JS. 2010. From sunlight to phytomass: on the potential efficiency of converting solar radiation to phyto-energy. *New Phytologist* **188,** 939-959.

Amthor JS. 1989. *Respiration and crop productivity*. Berlin: Springer-Verlag.

Angel J. 2009. The Water and Atmospheric Resources Monitoring Program. Illinois State Water Survey, University of Illinois at Champaign-Urbana. http://www.isws.illinois.edu.

Austin RB, Ford MA, Morgan CL. 1989. Genetic improvement in the yield of winter wheat: a further evaluation. *Journal of Agricultural Science* **112,** 295-301.

Bilgin DD, DeLucia EH, Clough SJ. 2009. A robust plant RNA isolation method suitable for Affymetric GeneChip analysis and quantitative real-time PCR. *Nature Protocols* **4,** 333-340.

Board JE, Harville BG. 1992. Explanations for greater light interception in narrow- vs. wide-row soybean. *Crop Science* **32,** 198-202.

Borras L, Slafer GA, Otegui ME. 2004. Seed dry weight response to source-sink manipulations in wheat, maize and soybean: a quantitative reappraisal. *Field Crops Research* **86,** 131-146.

Boyer JS. 1982. Plant productivity and environment. *Science* **218,** 443-448.

Busov VB, Brunner AM, Strauss SH. 2008. Genes for control of plant stature and form. *New Phytologist* **177,** 589-607.

Calderini DF, Dreccer MF, Slafer GA. 1997. Consequences of plant breeding on biomass growth, radiation interception, and radiation use efficiency in wheat. *Field Crops Research* **52,** 271-281.

Choe S, Dilkes BP, Fujioka S, Takatsuto S, Sakurai A, Feldmann KA. 1998. The *DWF4* gene of Arabidopsis encodes a cytochrome P450 that mediates multiple 22α-hydroxylation steps in brassinosteroid synthesis. *The Plant Cell* **10,** 231-243.

Choe S, Fujioka S, Noguchi T, Takatsuto S, Yoshida S, Feldmann KA. 2001 Overexpression of *DWARF4* in the brassinosteroid biosynthetic pathway results in increased vegetative growth and seed yield in *Arabidopsis*. *The Plant Journal* **26,** 573-582.

Clouse SD, Sasse JM. 1998. Brassinosteroids: essential regulators of plant growth and development. *Annual Review of Plant Physiology and Plant Molecular Biology* **49,** 427-451.

Condon AG, Richards RA, Rebetzke GJ, Farquhar GD. 2004. Breeding for high water-use efficiency. *Journal of Experimental Botany* **55,** 2447-2460.

Crafts-Brandner SJ, Salvucci ME. 2000. Rubisco activase constrains the photosynthetic potential of leaves at high temperature and CO_2. *Proceedings of the National Academy of Sciences of the United States of America* **97,** 13430-13435.

de Bossoreille de Ribou S, Douam F, Hamant O, Frohlich MW. 2013. Plant science and agricultural productivity: Why are we hitting the yield ceiling? *Plant Science* **210**, 159-176.

De Bruin JL, Pederson P. 2008. Yield improvement and stability for soybean cultivars with resistance to *Heterodera glycines* Ichinohe. *Agronomy Journal* **100**, 1354-1359.

Deprost D, Yao L, Sormani R, Moreau M, Leterreux G, Nicolai M, Bedu M, Robaglia C, Meyer C. 2007. The *Arabidopsis* TOR kinase links plant growth, yield, stress resistance and mRNA translation. *EMBO* **8,** 864-870.

Dermody O, Long SP, McConnaughay K, DeLucia EH. 2008. How do elevated CO_2 and O_3 affect the interception and utilization of radiation by a soybean canopy? *Global Change Biology* **14**, 556-564.

Driever SM, Lawson T, Andralojc PJ, Raines CA, Parry MAJ. 2014. Natural variation in photosynthetic capacity, growth, and yield in 64 field-grown wheat genotypes. *Journal of*

Experimental Botany **65,** 4959-4973.

Duvik DN, Cassman KG. 1999. Post-green revolution trends in yield potential of temperate maize in the north-central United States. *Crop Science* **39,** 1622-1630.

Etheridge DM, Steele LP, Langenfelds RL, Francey RJ, Barnola J-M, Morgan VI. 1996. Natural and anthropogenic changes in atmospheric CO_2 over the past 1000 years from air in Antarctic ice and firn. *Journal of Geophysical Research* **101,** 4115-4128.

Evans LT, Fischer RA. 1999. Yield potential: its definition, measurement, and significance. *Crop Science* **39**, 1544-1551.

Evans LT. 1993. *Crop evolution, adaptation, and yield*. Cambridge University Press, Cambridge.

Farquhar GD, Sharkey TD. 1982. Stomatal conductance and photosynthesis. *Annual Review of Plant Physiology and Plant Molecular Biology* **33,** 317-345.

Farquhar GD, von Caemmerer S, Berry JA. 1980. A biochemical model of photosynthetic CO_2 assimilation in leaves of C_3 species. *Planta* **149,** 78-90.

Fehr WR, Caviness CE, Burmood DT, Pennington JS. 1971. Stage of development descriptions for soybean, *Glycine max* (L.) Merr. *Crop Science* **11,** 929-931.

Fischer RA, Rees D, Sayre KD, Lu ZM, Condon AG, Larque-Saavedra A. 1998. Wheat yield progress associated with higher stomatal conductance and photosynthetic rate, and cooler canopies. *Crop Science* **38,** 1467-1475.

Food and Agricultural Organization of the United Nations. 2013. FAOSTAT Database on Agriculture. Rome, Italy.

Franceschi VR, Giaquinta RT. 1983. The paraveinal mesophyll of soybean leaves in relation to assimilate transfer and compartmentation. II. Structural, metabolic, and compartmental changes during reproductive growth. *Planta* **157,** 422-431.

Gifford RM, Thorne JH, Hitz WD, Giaquinta RT. 1984. Crop productivity and photoassimilate partitioning. *Science* **225**, 801-808.

Gilbert ME, Holbrook NM, Zwieniecki MA, Sadok W, Sinclair TR. 2011a. Field confirmation of genetic variation in soybean transpiration response to vapor pressure deficit and photosynthetic compensation. *Field Crops Research* **124,** 85-92.

Gilbert ME, Zwieniecki MA, Holbrook NM. 2011b. Independent variation in photosynthetic capacity and stomatal conductance leads to differences in intrinsic water use efficiency in 11 soybean genotypes before and during mild drought. *Journal of Experimental Botany* **62,** 2875-2887.

Gillespie KM, Xu F, Richter KT, McGrath JM, Markelz RJC, Ort DR, Leakey ADB, Ainsworth EA. 2012. Greater antioxidant and respiratory metabolism in field-grown soybean exposed to elevated O_3 under both ambient and elevated CO_2. *Plant, Cell and Environment* **35,** 169-184.

Godfray HC, Beddington JR, Crute IR, Haddad L, Lawrence D, Muir JF, Pretty J, Robinson S, Thomas SM, Toulmin C. 2010. Food security: the challenge of feeding 9 billion people. *Science* **327,** 812-818.

Gonzalez N, Beemster GTS, Inze D. 2009. David and Goliath: what can the tiny weed *Arabidopsis* teach us to improve biomass production in crops? *Current Opinion in Plant Biology* **12,** 157-164.

Guzman PS, Diers BW, Neece DJ, Martin SKS, Leroy AR, Grau CR, Hughes TJ, Nelson RL. 2007. QTL associated with yield in three backcross-derived populations of soybean. *Crop Science* **47,** 111-122.

Hay RKM. 1995. Harvest index: a review of its use in plant breeding and crop physiology. *Annals of Applied Biology* **126,** 197-216.

Holtan HE, Bandong S, Marion CM *et al.* 2011. BBX32, and Arabidopsis B-box protein, functions in light signaling by suppressing HY5-regulated gene expression and interacting with STH2/BBX21. *Plant Physiology* **156,** 2109-2123.

Hymowitz, T. 1970. On the domestication of soybean. *Economic Botany* **24,** 408-421.

Jin J, Liu K, Wang G, Mi L, Shen Z, Chen X, Herbert SJ. 2010. Agronomic and physiological contributions to the yield improvement of soybean cultivars released from 1950 to 2006 in Northeast China. *Field Crops Research* **115**, 116-123.

Koester RP, Skoneczka JA, Cary TR, Diers BW, Ainsworth EA. 2014. Historical gains in soybean (*Gylcine max* Merr.) seed yield are driven by linear increases in light interception, energy conversion, and partitioning efficiencies. *Journal of Experimental Botany* **65**, 3311-3321.

Kumudini S. 2002. Trials and tribulations: a review of the role of assimilate supply in soybean genetic yield improvement. *Field Crops Research* **75**, 211-22.

Leakey ADB, Xu F, Gillespie KM, McGrath JM, Ainsworth EA, Ort DR. 2009. Genomic basis for stimulated respiration by plants growing under elevated carbon dioxide. *Proceedings of the National Academy of Sciences of the United States of America* **106,** 3597-3602.

Libault M, Thibivilliers S, Bilgin DD, Radwan O, Benitez M, Clough SJ, Stacey, G. 2008. Identification of four soybean reference genes for gene expression normalization. *The*

Plant Genome **1,** 44-54.

Lobell DB, Cassman KG, Field CB. 2009. Crop yield gaps: their importance, magnitudes, and causes. *Annual Reviews of Environment and Resources* **34,** 179-204.

Lobell DB, Roberts MJ, Schlenker W, Braun N, Little BB, Rejesus RM, Hammer GL. 2014. Greater sensitivity to drought accompanies maize yield in the U.S. Midwest. *Science* **344,** 516-519.

Long SP, Bernacchi CJ. 2003. Gas exchange measurements, what can they tell us about the underlying limitations of photosynthesis? Procedures and sources of error. *Journal of Experimental Botany* **54,** 2393-2401.

Long SP, Zhu X-G, Naidu SL, Ort DR. 2006. Can improvement in photosynthesis increase crop yields? *Plant, Cell and Environment* **29,** 315-330.

Loomis RS, Amthor JS. 1999. Yield potential, plant assimilatory capacity, and metabolic efficiencies. *Crop Science* **39**, 1584-1596.

McDermitt DK, Loomis RS. 1981. Elemental composition of biomass and its relation to energy content, growth efficiency, and growth yield. *Annals of Botany* **48,** 275-290.

Miflin B. 2000. Crop improvement in the 21st century. *Journal of Experimental Botany* **51,** 1-8.

Monteith JL. 1972. Solar radiation and productivity in tropical ecosystems. *Journal of Applied Ecology* **9**, 747-766.

Monteith JL. 1977. Climate and the efficiency of crop production in Britain. *Philosophical Transactions of the Royal Society Series B-Biological Sciences* **281**, 277-294.

Morrison MJ, Voldeng HD, Cober ER. 1999. Physiological changes from 58 years of genetic improvement of short-season soybean cultivars in Canada. *Agronomy Journal* **91**, 685-689.

Morrison MJ, Voldeng HD, Cober ER. 2000. Agronomic changes from 58 years of genetic improvement of short-season soybean cultivars in Canada. *Agronomy Journal* **92**, 780-784.

Moss DN, Musgrave RB. 1971. Photosynthesis and crop production. *Advances in Agronomy* **23,** 317-336.

Mueller-Cajar O, Stotz M, Bracher A. 2014. Maintaining photosynthetic CO_2 fixation via protein remodeling: the Rubisco activases. *Photosynthesis Research* **119**, 191-201.

Murphy KA, Kuhle RA, Fischer AM, Anterola AM, Grimes HD. 2005. The functional status of paraveinal mesophyll vacuoles in response to altered metabolic conditions in soybean

leaves. *Functional Plant Biology* **32**, 335-344.

Nobel PS, Forseth IN, Long SP. 1993. Canopy structure and light interception. In *Photosynthesis and Production in a Changing Climate.* Chapman and Hall, London, pp 79-90.

Orf JH, Chase K, Adler FR, Mansur LM, Lark KG. 1999. Genetics of soybean agronomic traits: II. Interactions between quantitative trait loci in soybean. *Crop Science* **39,** 1652-1657.

Ort DR, Long SP. 2014. Limits on yields in the Corn Belt. *Science* **344,** 484-485.

Ort DR. 2001. When there is too much light. *Plant Physiology* **125,** 29-32.

Palomeque L, Liu LJ, Li WB, Hedges B, Cober ER, Rajcan I. 2009. QTL in mega-environments: I. Universal and specific seed yield QTL detected in a population derived from a cross of high-yielding adapted x high-yielding exotic soybean lines. *Theoretical and Applied Genetics* **119,** 417-427.

Parry MA, Reynolds M, Salvucci ME, Raines C, Andralojc PJ, Zhu X-G, Price GD, Condon AG, Furbank RT. 2011. Raising yield potential of wheat. II. Increasing photosynthetic capacity and efficiency. *Journal of Experimental Botany* **6**, 453-467**.**

Parry MAJ, Andralojc PJ, Scales JC, Salvucci ME, Carmo-Silva AE, Alonso H, Whitney SM. 2013. Rubisco activity and regulation as targets for crop improvement. *Journal of Experimental Botany* **64,** 717-730.

Parry MAJ, Flexa J, Medrano H. 2005. Prospects for crop production under drought: research priorities and future directions. *Annals of Applied Biology* **147,** 211-226.

Peng S, Laza RC, Visperas RM, Sanico AL, Cassman KG, Khush GS. 2000. Grain yield of rice cultivars and lines developed in the Philippines since 1966. *Crop Science* **40,** 307-314.

Phillips RL. 2010. Mobilizing science to break yield barriers. *Crop Science* **50,** S99-S108.

Population Division of the Department of Economic and Social Affairs of the United Nations Secretariat. 2011. *World Population Prospects: The 2010 Revision.* New York: United Nations.

Portis AR. 2003. Rubisco activase – Rubisco's catalytic chaperone. *Photosynthesis Research* **75,** 11-27.

Preuss SB, Meister R, Qingzhan X *et al.* 2012. Expression of the *Arabidopsis thaliana BBX32* gene in soybean increases grain yield. *PLoS ONE* **7,** e30717.

Raines CA. 2011. Increasing photosynthesis carbon assimilation in C3 plants to improve crop yield: current and future strategies. *Plant Physiology* **115,** 36-42.

Ray DK, Mueller ND, West PC, Foley JA. 2013. Yield trends are insufficient to double global crop production by 2050. *PLoS ONE* doi:10.1371/journal.pone.0066428.

Ray DK, Ramankutty N, Mueller ND, West PC, Foley JA. 2012. Recent patterns of crop yield growth and stagnation. *Nature Communications* **3**, 1293.

Rebetzke GJ, Condon AG, Richards RA, Farquhar GD. 2002. Selection for reduced carbon isotope discrimination increases aerial biomass and grain yield of rainfed bread wheat. *Crop Science* **42,** 739-745.

Reynolds M, Bonnet D, Chapman SC, Furbank RT, Manes Y, Mather DE, Parry MAJ. 2011. Raising yield potential of wheat. I. Overviews of a consortium approach and breeding strategies. *Journal of Experimental Biology* **62**, 439-452.

Reynolds MP, Rajaram S, Sayre KD. 1999. Physiological and genetics changes of irrigated wheat in the post-green revolution period and approaches for meeting projected global demand. *Crop Science* **39,** 1611-1621.

Reynolds MP, van Ginkel M, Ribaut J-M. 2000. Avenues for genetic modification of radiation use efficiency in wheat. *Journal of Experimental Biology* **51,** 459-473.

Richards RA. 2000. Selectable traits to increase crop photosynthesis and yield of grain crops. *Journal of Experimental Botany* **51,** 447-458.

Rincker K, Nelson R, Specht J, *et al.* 2014. Genetic improvement of US soybean in maturity groups II, II, and IV. *Crop Science* **54,** 1-14.

Rosegrant MW, Agcaoili M. 2010. *Global food demand, supply, and price prospects to 2010.* Washington, DC: International Food Policy Research Institute.

Ross-Ibarra J, Morrell PL, Gaut BS. 2007. Plant domestication, a unique opportunity to identify the genetic basis of adaption. *Proceedings of the National Academy of Sciences of the United States of America* **104,** 8641-8648.

Rowntree S, Suhre J, Specht J *et al.* 2014. Physiological and phenological responses of historical soybean cultivar releases to earlier planting. *Crop Science* **54,** 804-816.

Rowntree SC, Suhre JJ, Weidenbenner NH, *et al.* 2013. Genetic gain x management interactions in soybean: I. Planting Date. *Crop Science* **53**, 1128-1138.

Ruijter JM, Ramakers C, Hoogaars WM, Karlen Y, Bakker O, van den Hoff MJ, Moorman AF. 2009. Amplification efficiency: linking baseline and bias in analysis of quantitative PCR data. *Nucleic Acids Research* **37,** e45.

Sadras VO, Lawson C. 2011. Genetic gain in yield and associated changes in phenotype, trait plasticity and competitive ability of South Australian wheat varieties released between 1958 and 2007. *Crop Pasture Science* **62,** 533–549.

Sayre KD, Rajaram S, Fischer RA. 1997. Yield potential progress in short bread wheats in Northwest Mexico. *Crop Science* **37,** 36-42.

Severin AJ, Woody JL, Bolon Y-T, *et al*. 2010. RNA-Seq atlas of *Glycine max*: a guide to the soybean transcriptome. *BMC Plant Biology* **10**, 160.

Shearman VJ, Sylvester-Bradley R, Scott RK, Foulkes MJ. 2005. Physiological processes associated with wheat yield progress in the UK. *Crop Science* **45,** 175-185.

Sinclair TR, de Wit CT. 1975. Photosynthate and nitrogen requirements for seed production by various crops. *Science* **189,** 565-567.

Sinclair TR, Muchow RC. 1999. Radiation use efficiency. *Advances in Agronomy* **65**, 215–265.

Sinclair TR, Purcell LC, Sneller CH. 2004. Crop transformation and the challenge to increase yield potential. *TRENDS in Plant Science* **9,** 70-75.

Slafer GA, Andrade FH, Storre EH. 1990. Genetic improvement effects on pre-anthesis physiological attributes related to wheat grain yield. *Field Crops Research* **23,** 225-263.

Slattery RA, Ainsworth EA, Ort DR. 2013. A meta-analysis of responses of canopy photosynthetic conversion efficiency to environmental factors reveals major causes of yield gap. *Journal of Experimental Botany* **64,** 3723-3733.

Smith AM, Stitt M. 2007. Coordination of carbon supply and plant growth. *Plant, Cell and Environment* **30,** 1126-1149.

Specht JE, Diers BW, Nelson RL, Toledo JF, Torrion JA, Grassini P. 2014. *Yield gains in major U.S. field crops.* CSSA Special Publicantion 33. Eds. Smith S, Diers B, Specht J, Carver B. ASA, CSSA, and SSSA, Madison, WI, USA.

Specht JE, Hume DJ, Kumudinia SV. 1999. Soybean yield potential – a genetic and physiological perspective. *Crop Science* **39,** 1560-1570.

Specht JE, Williams JH. 1984. Contribution of genetic technology to soybean productivity – Retrospect and prospect. In: W.R. Fehr, editor, Genetic contribution to yield gains of five major crop plants. CSSA Special Publ. 7. CSSA and ASA, Madison, WI. p. 49-74.

Spokus K, Forcella F. 2009. Software tools for weed seed germination modeling. *Weed Science* **57,** 216-227.

Staswick PE, Zhang Z, Clemente TE, Specht JE. 2001. Efficient down-regulation of the major vegetative storage protein genes in transgenic soybean does not compromise plant productivity. *Plant Physiology* **127**, 1819-1826.

Stitt M, Lunn J, Usadel B. 2010. Arabidopsis and primary photosynthetic metabolism – more than the icing on the cake. *The Plant Journal* **61,** 1067-1091.

Stupar RM, Specht JE. 2013. Chapter four – insights from the soybean (*Glycine max* and *Glycine soja*) genome: past, present, and future. *Advances in Agronomy* **118,** 177-204.

Sulpice R, Pyl ET, Ishihara H *et al.* 2009. Starch as a major integrator in the regulation of plant growth. *Proceedings of the National Academy of Sciences of the United States of America* **106,** 10348-10353.

Tilman D, Balzer C, Hill J, Befort BL. 2011. Global food demand and the sustainable intensification of agriculture. *Proceedings of the National Academy of Sciences of the United States of America* **108**, 20260-20264.

Tollenaar M, Lee EA. 2002. Yield potential, yield stability, and stress tolerance in maize. *Field Crops Research* **75,** 161-169.

Twine TE, Kucharik CJ. 2009. Climate impacts on net primary productivity trends in natural and managed ecosystems of the central and eastern United States. *Agricultural and Forest Meteorology* **149,** 2143-2164.

Van Camp W. 2005. Yield enhancement genes: seeds for growth. *Current Opinion in Biotechnology* **16,** 147-153.

von Caemmerer S, Farquhar GD. 1981. Some relationships between the biochemistry of photosynthesis and the gas exchange of *Phaseolus vulgaris* cultivar Hawkesbury-wonder leaves. *Planta* **153,** 376-387.

Wilson RF. 2008. Soybean: market driven research needs. In: G. Stacey, editor, *Genetics and Genomics of soybean.* Spring Science + Business Media, New York. 3-15.

Woodward FI. 1987. Stomatal numbers are sensitive to increases in CO_2 from pre-industrial levels. *Nature* **327,** 617-618.

Wullschleger S, Loewith R, Hall MN. TOR signaling in growth and metabolism. *Cell* **124,** 471-484.

Yendrek CR, Ainsworth EA, Thimmapuram J. 2012. The bench scientist's guide to statistical analysis of RNA-Seq data. *BMC Research Notes* **5,** 506.

Yin Z, Meng F, Song H, Wang X, Xu X, Yu D. 2010. Expression quantitative trait loci analysis of two genes encoding rubisco activase in soybean. *Plant Physiology* **152,** 1625-1637.

Zheng TC, Zhang XK, Yin GH, Wang LN, Han YI, Chen L, Huang F, Tang JW, Xia XC, He ZH. 2011. Genetic gains in grain yield, net photosynthesis and stomatal conductance achieved in Henan Province of China between 1981 and 2008. *Field Crops Research* **122,** 225-233.

Zhu X-G, Long SP, Ort DR. 2008. What is the maximum efficiency with which photosynthesis can convert solar energy into biomass. *Current Opinion in Biotechnology* **19,** 153-159.

Zhu X-G, Long SP, Ort DR. 2010. Improving photosynthetic efficiency for greater yield. *Annual Review of Plant Biology* **61**, 235-261.